'Mental rotation', pictured rotation, and tandem rotation in depth

Keith K. Niall *

Aircrew Training Research Division, Human Resources Directorate, Armstrong Laboratory (AL / HRA), 6001 South Power Road, Bldg. 558, Mesa, AZ 85206-0904, USA

Abstract

The mental rotation effect in depth is qualitatively different from mental rotation in the picture plane. The magnitude of angular difference in depth that is depicted between two static shapes has been thought to predict mean response time for same–different comparisons on shapes in the mental rotation effect. The *tandem rotation effect* provides a counterexample to that hypothesis. Two planar shapes are depicted as separated by a small and fixed angular difference in depth; the pair of shapes is then depicted as tilted in depth. Mean response time to compare these shapes varies nearly linearly with the magnitude of the yoked rotation – though angular difference is held constant. The slope of this response time function is very close to that for single rotation in depth. The tandem effect supports a claim that mean response time varies as a function of the change in area of planar shapes that are depicted to rotate away from the picture plane, rather than as a function of the angular difference of one shape from another. The tandem rotation effect is not found to obtain for rotations in the picture plane. A conclusion drawn from these and other results is that an hypothesis of mental rotation is neither necessary nor sufficient to explain changes in response times for the simultaneous comparison of planar shapes pictured in depth. A piecewise continuous trigonometric function is proposed to describe response times for the comparison of planar shapes that are depicted to rotate in depth. The characteristic Shepard–Metzler response time function for complex solids in depth is shown to be a definite integral of that trigonometric function.

---

* E-mail: niall@hrlban1.aircrew.asu.edu.

Report Documentation Page

1. REPORT DATE **JAN 1997**	2. REPORT TYPE **Final**	3. DATES COVERED **01-01-1995 to 01-11-1996**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
'Mental Rotation', Pictured Rotation, and Tandem Rotation in Depth	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER **62202F**

6. AUTHOR(S)	5d. PROJECT NUMBER **1123**
Keith Niall	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Armstrong Laboratory/HRA,Aircrew Training Research Division,6030 South Kent Street,Mesa,AZ,85212-6061**	8. PERFORMING ORGANIZATION REPORT NUMBER **AL; AL/HRA**

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061**	10. SPONSOR/MONITOR'S ACRONYM(S) **AFRL; AFRL/RHA**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AFRL-RH-AZ-JA-1997-**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Published in Acta Psychologica, 95, 31-83 (1997)

14. ABSTRACT

The mental rotation effect in depth is qualitatively different from mental rotation in the picture plane. The magnitude of angular difference in depth that is depicted between two static shapes has been thought to predict mean response time for same-different comparisons on shapes in the mental rotation effect. The tandem rotation effect provides a counterexample to that hypothesis. Two planar shapes are depicted as separated by a small and fixed angular difference in depth; the pair of shapes is then depicted as tilted in depth. Mean response time to compare these shapes varies nearly linearly with the magnitude of the yoked rotation - though angular difference is held constant. The slope of this response time function is very close to that for single rotation in depth. The tandem effect supports a claim that mean response time varies as a function of the change in area of planar shapes that are depicted to rotate away from the picture plane, rather than as a function of the angular difference of one shape from another. The tandem rotation effect is not found to obtain for rotations in the picture plane. A conclusion drawn from these and other results is that an hypothesis of mental rotation is neither necessary nor sufficient to explain changes in response times for the simultaneous comparison of planar shapes pictured in depth. A piecewise continuous trigonometric function is proposed to describe response times for the comparison of planar shapes that are depicted to rotate in depth. The characteristic Shepard-Metzler response time function for complex solids in depth is shown to be a definite integral of that trigonometric function.

15. SUBJECT TERMS
Mental rotation; Pictured rotation; Tandem rotation; Static shapes; Shepard-Metzler response time function; Visual perception; Linear perspective; Spatial imagery

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Public Release**	**52**	

1. Introduction

The visual perception of pictures and drawings resembles the visual perception of nearby objects in that pictures and drawings reproduce a part of the conditions under which light arrives at the eyes. The perception of pictures and drawings most resembles the perception of solids as one stands still and looks at nearby solids. This resemblance is easily overstated; an observer's estimation of shape from pictures is based on convention in ways that the observer's estimation of the shape of nearby objects is not. It is a mistake to suppose that the two situations are to be explained in just the same way. It is also a mistake to suppose that the perception of a rotation which is pictured to take place in depth is to be explained in the same way as the perception of a rotation which takes place in the picture plane. These two conditions have been confounded in interpretation of what is called 'the mental rotation effect'.

The mental rotation effect concerns the time it takes an observer to respond 'same' or 'different' to a pair of shapes, where the shapes differ in angular orientation, and may differ in handedness (or in other features). Over numbers of trials, an observer's response times can be correlated with magnitudes of difference in angular orientation between the shapes. 'Angular orientation' has been taken to denote either angular orientation in the picture plane or pictured orientation in depth, indifferently. The discoverers of the mental rotation effect (R. Shepard and J. Metzler, 1971) found response times to be a regular and roughly linear function of difference in angular orientation. They believe that the effect is evidence for a mental process of rotation. (The mental rotation effect will often be referred to as 'the Shepard–Metzler effect' in what follows. Similarly, Cohen and Kubovy (1993, p. 352) say: "the term 'mental rotation' ... is ambiguous. It sometimes refers to a *mental operation* that involves an imagined reorientation of form, and sometimes refers to a *task* assumed to require the mental operation.")

Mental rotation is part of an explanation of perceptual phenomena by *natural geometry*. Of course observers may compare shapes with ruler and protractor or with theodolite and tacheometer, but this is not what is called natural geometry. Instead geometry is thought to be 'built into' psychological processes of visual perception, as though the eyes themselves could apply methods of geometry – as though the eye were a theodolite for the visual field. Such an explanation by natural geometry conflates visual comparison with physical measurement. This application of geometry is called natural in that it is supposed to be innate: "... the system of constraints that governs the projections and transformations of such bodies in space must long ago have become internalized as a powerful, though largely unconscious, part of our perceptual machinery" (Shepard, 1978, p. 136). This makes the theory of vision parasitic on geometry: it is unclear what could be meant by a 'mental operation of rotation', except by reference to physical operations of rotation. Mental rotation is a thesis both about vision and about visual imagery, but as Pylyshyn (1981, pp. 199–200) remarks:

"... there is only one empirical hypothesis responsible for the predictive success of the whole range of imagistic models and ... nearly everything else about such models consists of free empirical parameters added ad hoc to accomodate particular

experimental results. The one empirical hypothesis is just this: *When people imagine a scene or an event, what goes on in their minds is in many ways similar to what goes on when they observe the corresponding event actually happening."*

At first mental rotation was presented as a simple operation, a mental analogue of the physical operation of rotation in space. Since then the story of mental rotation has become far more complicated. Does this mean that the description of a simple operation has been embellished by detail, or could it mean that the Shepard–Metzler effect never did stem from the mental analogue of a simple operation? There are a couple of ways that the Shepard–Metzler effect might prove to be something other than the observable phenomenon of an analogue of rotation hidden in the mind. One is that the very notion of a mental process may be incoherent if it is predicated on similarity to physical processes, and another is that the Shepard–Metzler effect may fail to conform to basic facts about the physical operation of rotation. Peter Hacker (1990, p. 307) takes the former philosophical approach when he says:

"But when doing philosophy, one is prone to say 'Thinking is a mental activity', in order to distinguish it from *physical* activities. This *is* misleading... Similarly, saying that thinking is a mental activity typically leads philosophers and psychologists into futile investigations about the 'materials' of thought (words, images, internal representations) and about mental operations allegedly constitutive of thinking (e.g. the psychologist's supposition that thinking about whether two drawings are of one and the same object at different orientations involves rotating images in mental space at constant velocity). ... It induces the wrong pictures of thinking and generates misleading questions. It is, therefore, a move best avoided."

If Hacker is correct that current psychological investigation into mental rotation is founded in conceptual confusion, then it is pointless to speak of a sort of rotation as being mental as opposed to physical. But the other psychological approach may be more convincing to psychologists: perhaps the Shepard–Metzler effect does not conform to basic facts about physical operations of rotation.

Four empirical claims have been proposed on the hypothesis that mental rotation bears a simple relation to response time, as speed of rotation does with respect to time in kinematics. The first is that response times to decisions of 'same' or 'different' increase linearly on average as the angular difference between two shapes increases. (cf. R. Shepard and J. Metzler, 1971, p. 701; Cooper and Shepard, 1973, p. 86. Note that the term 'angular difference' is ambiguous with regard to the dimensionality of the shapes.) The second claim is that response times continue to increase as far as an angular difference of 180°, past which response times decrease at the same rate to 360°. The third claim is that this rate – the slope of the function of angular difference versus response times – is identical for rotation in the picture plane and for perspective pictures of rotation in depth (Metzler and Shepard in Shepard and Cooper, 1982, pp. 44–45 and 50; Shepard and Judd, 1976, p. 954). The fourth is that, for pictured rotations in depth, the degree of similarity between the perspective traces of the two shapes in the picture plane has no independent effect on response times (cf. Metzler and Shepard, 1974, p.

168). In the final analysis, none of these claims withstands empirical scrutiny for planar shapes depicted to rotate in depth.

Shepard mentions another possibility: another explanation can be given for those response time results. He says (in Shepard and Cooper, 1982, p. 116): "it may be that someone will be able to formulate a theory that satisfyingly accounts for this particular set of facts without invoking any such concepts as mental imagery or mental rotation". The alternate explanation supposes that observers apprehend certain geometrical *invariants*: geometrical properties that are preserved as light travels from reflective surfaces to the eyes. Shepard's (Metzler and Shepard in Shepard and Cooper, 1982, p. 64) name for this type of explanation is the *comparison of rotationally invariant structural codes*. He describes how such an explanation might proceed, but does not defend this explanation as vigorously as his own. Shepard (Ibid.) reifies the alternate explanation in a psychological process, as follows:

"... the subject detects the presence and interrelationships of the basic components of one of the two-dimensional drawings – particularly, the variously oriented straight lines, the several types of vertices by which they are connected and, presumably, something of the structural relationships among these components within the two-dimensional pattern. Then, on the basis of some higher-level processing of these extracted features and their interrelationships, an internal representation, code, or verbal description is generated for each picture separately that captures the intrinsic structure of the three-dimensional object in a form that is independent of the particular orientation in which that object happens to be displayed."

He has made his description general, so he can address the general claims characteristic of this kind of explanation. Some of these claims are that: (1) observers are able to appreciate certain geometrical properties of drawings; (2) the relevant geometrical properties are properties found in plane geometry; and (3) these properties are independent of changes in an object's spatial orientation, and of subsequent perspective projection in a drawing. (Simply put, Shepard's intended properties are invariants of projective geometry.) Then he describes what seem to be problems or concerns with the approach. These are some of the response time results just mentioned:

Result 1.   Response times increase monotonically with the angular difference in orientation of the objects depicted in drawings.

Result 2.   The rate of this increase in response times is equally great for rotations in the picture plane as for depicted rotations in depth.

Result 3.   The increase in response times is very nearly linear for angular differences up to 180°.

The first result may seem incompatible with an account based on invariants, because the relevant geometric properties are independent of orientation. If the measure of those properties does not covary with angular difference, how could they account for a positive relation between angular difference and response time? This difficulty depends on the claim that observers are able to detect or compute these geometrical properties of

drawings: that observers are able to perceive the magnitudes of those properties, or that they are acquainted with their magnitudes through a process of 'unconscious inference'. If observers were able to perceive these magnitudes exactly, their ability would explain nothing of the relation between angular difference and response time: the magnitudes of the geometric properties in question are unrelated to angular difference (the issue has been raised by Jolicoeur, 1990, pp. 396–397, among others). But why should one expect observers to perceive the exact measure of these properties? Estimates of many other visible properties are related to their physical magnitudes by nontrivial psychophysical functions, and those estimates are affected by variations in conditions of observation. Perhaps the perception of invariants is not exact and unbiased either; it could be affected by variation in other conditions – such as pictured differences in orientation. Although the value of a geometric invariant does not serve to explain a relation between orientations and response times, changes in the discriminability of invariants could serve to explain the relation. This would answer Shepard's first objection to the 'comparison of rotationally invariant structural codes'.

Shepard's second objection to an account based on invariants depends on his second result: that the rate of increase in response times with angular difference is equally great for angular differences in the picture plane as for angular differences depicted in depth. Metzler and Shepard (in Shepard and Cooper, 1982, p. 66) say:

> "Nor do such theories provide a ready account for the equivalence of the slopes of the reaction-time functions for the picture-plane and depth pairs. For, in order to explain the dependence of reaction time on angular difference, we must suppose that the features that are being compared are the features of the two-dimensional drawings, which differ more and more with angular departure, and not the features of the three-dimensional objects, which are the same regardless of orientation."

This initial claim that response time functions are identical for the two cases might be revised if the slope of the response time function for rotation in depth were found to be much greater than the slope of the response time function for rotation in the picture plane. Such differences in slope could be taken to indicate a difference in the 'rate of mental rotation'. Yet such a finding could be re-interpreted in several ways. Differences due to dimensionality between the mental rotation effect in depth and the mental rotation effect in the picture plane can be attributed to complexity of shape instead, or to other factors. Shepard and Cooper (1982, p. 178) say:

> "...most studies employing three-dimensional objects as stimuli have used simultaneous presentation whereas most studies employing two-dimensional objects have used comparison of a single visual stimulus with a memory presentation. We suspect that it is this procedural difference rather than the difference in dimensionality that is the principal determiner of rate of mental rotation."

Not only the method of presentation, but also the classification of stimuli has been used to explain apparent differences between the Shepard–Metzler effect in the plane and the Shepard–Metzler effect in depth. Shepard (cited in S. Shepard and D. Metzler, 1988, p. 4) claims that all objects are represented in three-dimensional space, in effect that all pictorial stimuli are three-dimensional. Then he interprets the difference between rotation in the plane and depicted rotation in depth as a difference between rotating a

thin plate (i.e., that which is visible only as a thin plate) and rotating a solid form. He attributes differences in response times between the two conditions to this difference in shape, rather than to the effect of dimensionality. Consequently differences in dimensionality cannot affect response times, since by this definition there are no differences in dimensionality.

Differences between rotation in the picture plane and depicted rotation in depth might still be construed as the effect of a change in the axis of rotation. The slope of the response time function does differ over several directions of the axis of depicted rotation (Parsons, 1987, p. 49). Lawrence Parsons claims that both the axes of rotation and the intrinsic geometry of the stimulus shapes will influence response times (see also Parsons, 1987a,1995). But Shepard conflates these effects with the effect of shape, as he conflated dimensionality with shape before. He interprets differences in response time that are occasioned by direction of rotation as differences due to relative direction of rotation, where direction of rotation is relative to the 'natural' axes of symmetry of the stimulus shape. These are axes of reflection or rotation that leave the stimulus shape unchanged, or which leave a component of the shape unchanged. Carlton and Shepard (1990, p. 181) say that "... the psychological accessibility of the transformational axes prescribed by kinematic geometry depends on the alignment of those axes with the natural axes of the environmental frame and, perhaps particularly, with the inherent axes of symmetry of the object itself''.

By now it ought to be clear that the empirical validation of this basic claim about the mental rotation effect is not a straightforward matter. New results have been assimilated by making response times depend on shape as well as angular difference, and by altering what counts as a relevant description of shape. Whether it be right or wrong, the claim that the rate of increase in response times is equally great for rotations in the picture plane as for depicted rotations in depth is a claim that has been rendered indefeasible by plain demonstration. Therefore it is no simple objection to an account based on invariants, before an account based on invariants has been extended to cover a variety of stimulus shapes and forms. Another attempt to show a difference in response times between plane rotation and rotation in depth will be included in the experiments that are reported in this article.

Shepard's third objection depends on his result that increase in response times with angular difference is very nearly linear for angular differences up to 180 degrees. Yet this pattern of response times differs between shapes. Pierre Jolicoeur and his co-workers (Jolicoeur et al., 1985, p. 125) report that response times increase linearly with the rotation of planar shapes in the picture plane, while they increase piecewise linearly (or non-linearly, one can surmise) in the picture plane rotation of shapes depicted as solid. (There are also differences between relatively flat and solid shapes that are depicted to rotate in depth: see Bauer and Jolicoeur, 1996). (Actually response times to pairs of planar shapes increase up to 90° of angular difference from the picture plane in depth, and then decrease to 180°, as we shall see. Fig. 1 provides a guide to the imagination.) So none of Shepard's problems or concerns with an explanation based on the perception of geometric invariants are telling problems: those objections can be answered.

The reasons for and the implications of a theory based on invariants have been construed several different ways in explanation of the Shepard–Metzler effect. As an
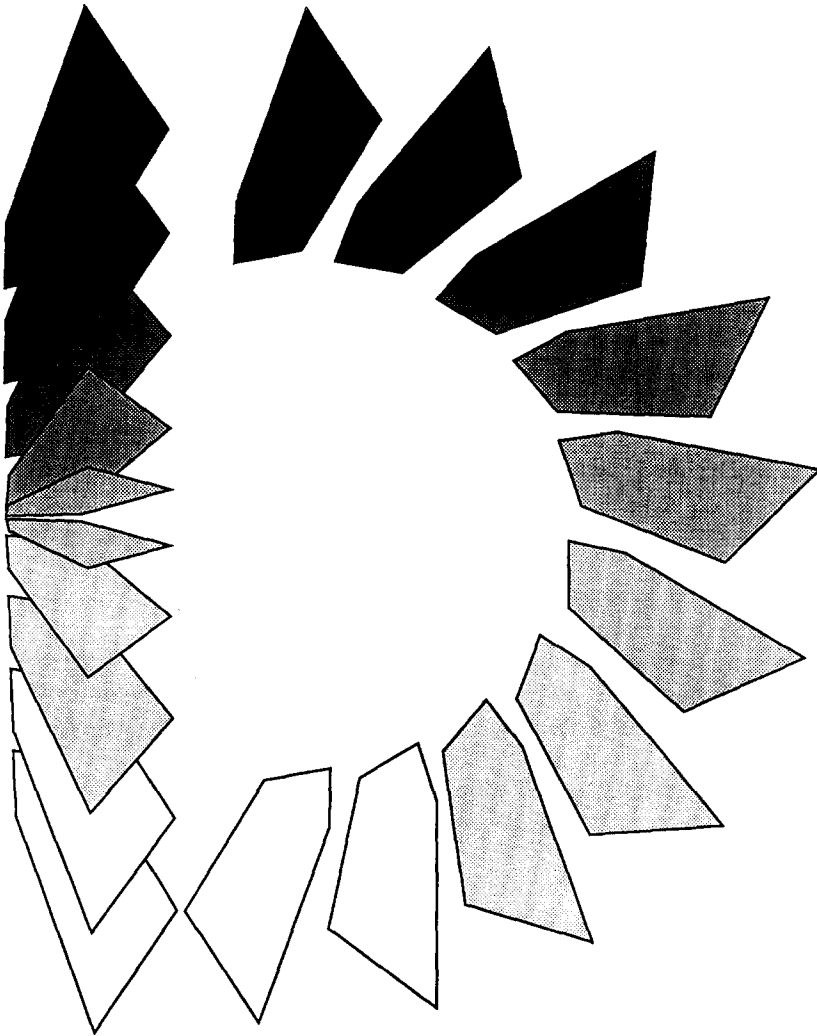
Fig. 1. A rotation in the picture plane (illustrated by the arc of shapes) consists of a simple change in the orientation of a shape, while a 'rotation in depth' (illustrated by the column of overlapping shapes) consists of projective transformations of the shape, which serve to depict changes in orientation about an axis horizontal to the page.

example, Van Gool et al. (1994, p. 559) hypothesize: "Recognition times in perception sometimes vary linearly with some pose parameter. Examples are mental rotation and scaling ... These results are not necessarily in contradiction to the use of invariants. A possible explanation of these effects would be the time needed to find the corresponding reference points used in the extraction of the invariants." Sometimes home truths about the application of geometric invariants and geometric transformations are mixed with narrow psychological assumptions. If, say, the discriminability of invariants decreases

monotonically with pictured rotation in depth from the picture plane, then there is a basis in the *perception* of invariants to explain the mental rotation effect.

The example serves to emphasize that a tally of geometric invariants is not the same as a psychological theory of the perception of those invariants and transformations. At times that difference has been masked by injudicious use of language. There is an ambiguity to the phrase 'a theory of vision based on invariants', that can mask the difference between an algebra of invariants as geometric properties, and an account of the perception of invariants. If one finds a geometric property that is perceived accurately and unequivocally, one might call it an invariant, without bothering to name the set of transformations under which it is an invariant, or without bothering to name the geometry which is specified by the invariant of interest. Yet not every property that is perceived counts as an invariant, if 'invariant' is to retain any geometric meaning. An invariant is an invariant under a set of transformations, and any geometric property is an invariant under some set of transformations, though it may not be a 'basic' invariant. Invariants under linear perspective are projective invariants (or affine invariants, if nearby objects are not under consideration). The importance of calling a property 'an invariant' lies in the property's context of transformations, and its place in geometry. Psychological considerations do not change what it means for a property to be an invariant, and this is of prime importance for any psychological account of the perception of invariants. The measurement of invariants in vision research should begin with optical considerations, not with psychological considerations. (As an example, Foster and Simmons (1994, p. 45) approach the 'problem of human object recognition' in this straightforward way. See Appendix A for more on the related 'problem of visual space'.)

Our theoretical imagination can outrun our understanding when we take an uncomplicated change in discriminability as evidence of a complicated mental realm. ("It is amazing how hard it is to get back to the idea that we do, after all, normally perceive what is out there, not something 'in here'"; Putnam, 1990, p. 251.) In due time it is the truth of predictions about response times rather than their source that will concern us; the inspiration or the original motivation for a psychological theory assumes less importance. So it may be for the Shepard–Metzler effect in depth: one day the effect may be recognized as significant, without being recognized as evidence for a mental process of rotation on analogy with rotation in kinematics. It may indeed be premature "to propose that spatial imagination has evolved as a reflection of the physics and geometry of the external world" (Cooper and Shepard, 1984, p. 114). My aim is to change the interpretation of the Shepard–Metzler effect in depth, at least for one broad class of shapes, by supplanting that interpretation with a description that makes no appeal to the same model of physical processes. Contrary to Shepard's fourth claim about the Shepard–Metzler effect in depth, variation in mean response time can be described in terms of the perspective traces of two shapes in the picture plane. This alternate description relates response times to geometric invariants and geometric transformations, rather than to magnitudes of angular difference. The geometric invariant is a ratio of areas, that stands for the degree of compression of a shape that is pictured to lie at a slant in depth. This alternate description of the mental rotation effect in depth can be extended, to predict a novel effect, a *tandem* rotation effect in depth.

The new description does not involve rotation; all that is needed can be found in the picture plane, right in front of one's nose.

2. Experiment 1: Rotation in the picture plane and rotation in depth

There is a difference that sometimes is acknowledged between the pattern of response times for rotation in depth and the pattern of response times for rotation in the picture plane. (At other times this difference is attributed to the effect of a difference in shape or dimensionality between stimuli in the two conditions: see Appendix B.) The difference is in the slope of the regression line for the response time function over magnitude of angular difference. That slope is greater (i.e., steeper) for rotation in depth than it is for rotation in the picture plane (cf. Parsons, 1987). Under the hypothesis of mental rotation, this can be interpreted as a change in the rate of mental rotation; mental rotation in depth proceeds less quickly than mental rotation in the plane.

This experiment is meant to demonstrate those response time differences that may exist between rotation in depth and rotation in the picture plane, for planar shapes. A second question is addressed by this demonstration as well. Mental rotation is thought to be a process by which shapes are compared indifferently of perspective effects. It is meant to be a symmetric process; the angular difference that is formed by rotating one shape of an unlike pair is equivalent to the angular difference of the same magnitude that is formed by rotating the other member of the pair in place of the first. If response times are not symmetric in the way angular differences are, then perhaps those response times do not reflect a process of comparison over angular difference at all. Instead they may reflect a perspective effect that applies to shapes one at a time. Another way to express this is to say that response times may actually reflect differences in the plane geometry of the display, instead of reflecting abstract operations in an imagined three-dimensional space.

2.1. Method

2.1.1. Subjects
Eight students and employees of DCIEM (the Defence and Civil Institute of Environmental Medicine, North York, Ontario, Canada) were tested, of whom four were men and four were women. All the observers had normal uncorrected acuity. The mean age of the observers was 26.0 years. Each subject signed a document indicating informed consent in the study, as did all subjects in the other experiments that will be reported.

2.1.2. Stimuli
The shapes were presented on the graphics console of an IRIS GT microcomputer. The aspect ratio of the display was chosen to ensure correct perspective projection. Two planar pentagons were presented on each trial; they were green in colour on a black background, with full simulation of lighting effects. The reflectance of these shapes was moderately specular, as for a dull metallic surface. The ambient illumination of the

scene was white (the background remained black), while local illumination was blue-green (*RGB* 0.5 : 1.0 : 1.0). Two local light sources of equal brightness were simulated. The light sources do not appear in any display themselves. The relative arrangement of these light sources and the viewpoint of the scene can be given in terms of a system of three-dimensional (X, Y, Z) coordinates. Before rotation, the shapes are conceived to lie in a coronal plane, the xy plane. The viewpoint distance for the scene is given a value of 10 from the origin along the z-axis. Naturally the viewpoint lies away from the screen, towards the observer. The (X, Y, Z) coordinates of the local light sources are: (2.5, 2.5, 30) and (-2.5, -2.5, 30). The rendering of the lighting effects was complete before the beginning of each trial.

Each day of testing consisted of many trials, that is, many pairs of shapes. The observers used a three-button 'mouse' device to register their responses. A small circular green spot appeared for three seconds between trials, to provide a fixation point. A short tone announced the beginning of each new trial. The sequence of trials paused when the observer pressed the middle button at the same time as the green spot was present. The sequence resumed when the middle button was pressed a second time. The observer's task was to respond 'same' or 'different' to each pair of shapes. 'Different' meant either that the shapes differed in handedness, or that they were wholly different shapes. Shapes that were wholly different were projectively incongruent. The left and right buttons of the mouse device registered responses of 'same' or 'different' (depending on the experimental condition). The buttons were labelled 'same' or 'different' as appropriate. Each trial was timed from onset to response; a response of 'same' or 'different' stopped the timer and ended the trial. Trials could last no longer than 20 seconds. Trials were re-scheduled if no response was registered in less than that time. Response times and decisions were tallied automatically, together with independent variables for accuracy of recordkeeping. Usually the entire session lasted less than an hour, though the time that was required to complete a sequence of trials varied from day to day. These aspects of the test situation were retained for the other experiments that will be reported (apart from a change in viewpoint in Experiment 3).

Three pairs of shapes (see Fig. 2) were presented in single rotation. They were Shapes *1* and *4*, Shapes *2* and *4*, and Shapes *3* and *4*.

These *Pairs* were presented in four arrangements or *Orders*. If the shapes are {*a*, *b*}, then the orders are (*a*, *a*), (*b*, *b*), (*a*, *b*), and (*b*, *a*). One of the shapes in each pair was either rotated in the picture plane (*z-Axis* rotation), or pictured to rotate in depth (*x*-axis rotation). These shapes were rotated at one of seven *Orientations* to upright: 20°, 40°, 60°, 80°, 100°, 120°, and 140°. The 'centre of mass' of the shapes did not translate with

Fig. 2. The stimuli for Experiment 1 consist of three pairs of shapes. A single shape appears as a member of each pair. The relative coordinates of the shapes (in reading order) are:

Shape 1:   [(9.76, 5.63), (7.76, 7.49), (6.81, 4.38), (3.92, 1.39), (8.47, 0.76)]
Shape 4:   [(6.97, 9.12), (4.44, 2.17), (4.38, 0.03), (6.77, 0.47), (9.64, 5.17)]
Shape 2:   [(4.60, 0.97), (4.70, 2.82), (4.31, 4.12), (1.44, 6.75), (1.28, 0.57)]
Shape 4:   [(6.97, 9.12), (4.44, 2.17), (4.38, 0.03), (6.77, 0.47), (9.64, 5.17)]
Shape 3:   [(4.52, 4.88), (5.00, 3.06), (6.93, 6.09), (8.95, 9.56), (5.20, 6.08)]
Shape 4:   [(6.97, 9.12), (4.44, 2.17), (4.38, 0.03), (6.77, 0.47), (9.64, 5.17)]
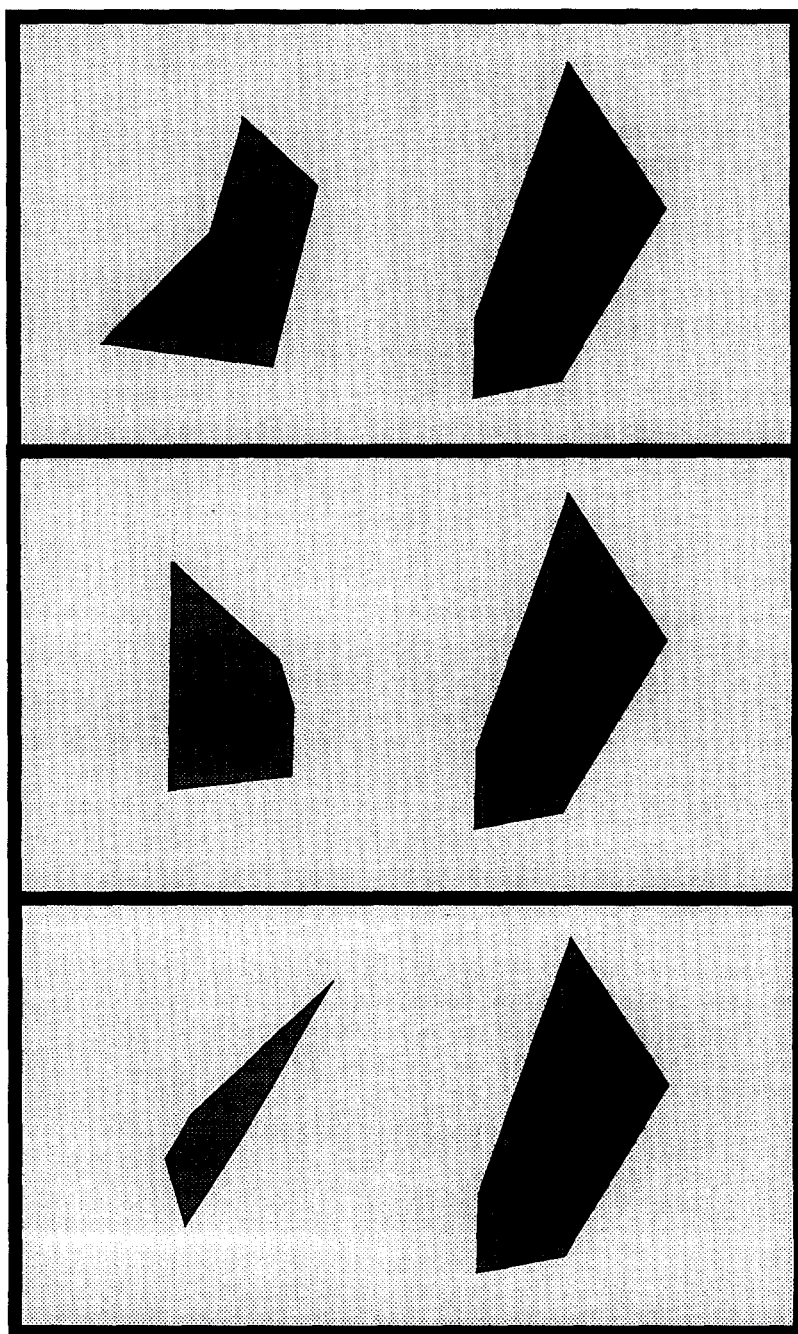
rotation of the shapes. Observers saw 56 distinct pairs of shapes on each day of testing. These 56 trials were randomized as a block of trials. Observers saw four *Blocks* each day; trials within each block were randomized independently.

Different pairs of shapes were presented on separate days of testing. The rotated shape appeared on the left or the right *Side* of the display on separate days, as well. The order of these conditions was randomized over days for each observer. The experiment has a six-way $(4 \times 2 \times 7 \times 4 \times 3 \times 2$, that is, *Order × Axis × Orientation × Block × Pair × Side*) repeated-measures design. There were a total of 1344 trials over six days for each observer in the experiment.

2.2. Results

The dependent measure for the main analysis, as for those of all subsequent experiments, is the natural logarithm of response time. (All further references to 'logarithm', 'ln', or 'log' should be taken to signify natural logarithms.) The principal motivation for applying the logarithmic transformation is the treatment of long outlying response times. The effect of the transformation is not gross over the range in question; its effect is that the distribution of response times becomes less skewed, and the tails of the distribution become more symmetric (Ratcliff, 1993). Times are included for all responses: that is, no 'trimming' of outlying response times has been applied over and above the logarithmic transformation. And unless otherwise stated, response times are included both for correct responses and for trials on which an error occurred.

A six-way analysis of variance showed significant effects of *Axis* $(F(1,7) = 9.55, p \leq 0.05)$ and *Orientation* $(F(6,42) = 28.92, p \leq 0.01)$ on the dependent measure of log response time, for conditions in which 'same' was the appropriate response. There was also a significant effect of *Pair × Side* $(F(2,14) = 7.49, p \leq 0.01)$. This indicates that response times depend on the *particular* shape that is rotated in depth, though the interaction could be confounded with the effect of practice over days. The Greenhouse–Geisser correction was applied to the degrees of freedom used to obtain the critical F statistics for these comparisons (Greenhouse and Geisser, 1959). This conservative procedure will be applied in all further significance levels reported for analyses of variance. A six-way analysis of variance was also performed on the conditions for which 'different' was the appropriate response. Again there were significant effects of *Axis* $(F(1,7) = 8.65, p \leq 0.05)$ and *Orientation* $(F(6,42) = 10.10, p \leq 0.05)$. There was a significant interaction of these two factors, *Axis × Orientation* $(F(6,42) = 7.17, p \leq 0.05)$. The particular shape that appears as a comparison in a *Pair* has a significant effect $(F(2,14) = 6.72, p \leq 0.05)$, as does the member of a pair that is rotated $(Pair × Side: F(2,14) = 15.20, p \leq 0.01)$. The latter findings are consonant with the dependence of response times on stimulus differences found by Petrusic et al. (1978, p. 142), who say: "The present finding of stimulus-dependent rotation speed is troublesome for the conventional view of the process of mental rotation."

The interaction of *Axis × Orientation* was not significant for the conditions in which 'same' was the appropriate response $(F(6,42) = 5.19, \text{NS})$. Certainly a stringent criterion has been applied, but an interesting reason can be given for the nonsignificance of this interaction term. A difference in the slope of the mental rotation function between
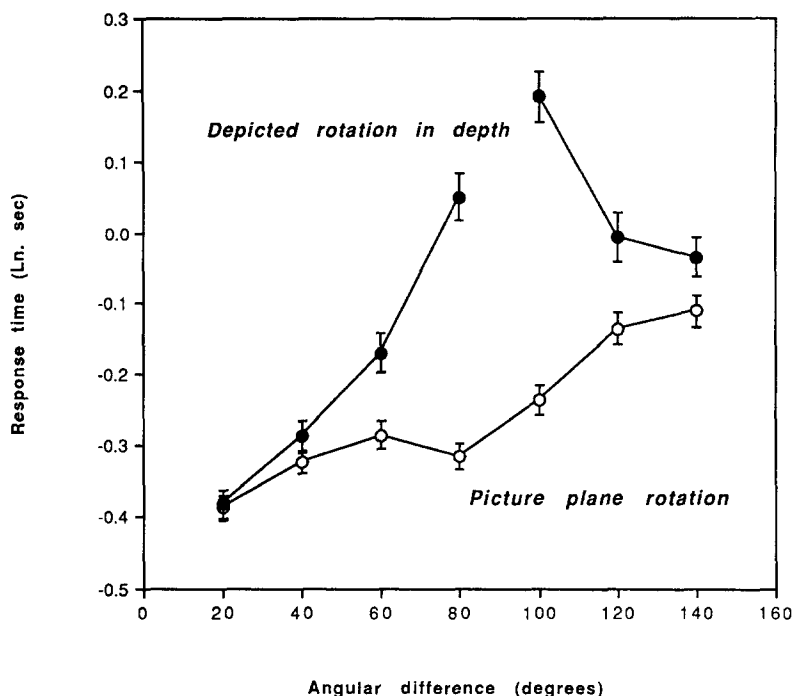
Fig. 3. Response times for depicted rotation in depth display a different pattern than do response times for rotation in the picture plane in Experiment 1. Mean response times (in ln seconds) are plotted against angular difference (in degrees); standard error bars are included. Mean response times for 'same' responses increase regularly with rotation in the picture plane (open circles), while mean response times for 'same' responses increase regularly only to 90° of angular difference with rotation in depth (filled circles). Only correct responses of 'same' were included for these calculations; each mean is based on a maximum of 384 trials. The results foreshadow the model that will be presented in Fig. 15, and they parallel the results presented in Fig. 16.

rotation in the picture plane and depicted rotation in depth will be reflected in a significant *Axis × Orientation* interaction, given that response times increase regularly over the range of angular difference. Response times are expected to increase monotonically from 0° to 180° of angular difference, yet they do not for the depicted rotation of planar shapes in depth (see Fig. 3). Here the mean response time for an angular difference of 140° is appreciably less than the mean response time for an angular difference of 100°, both for 'same' responses to rotation in depth, and for 'different' responses to rotation in depth. This incidental finding suggests that the maximum value of the mental rotation function for planar shapes lies near 90° of angular difference from the picture plane, not at 180° as for rotation within the picture plane.

2.3. Discussion

There is a difference in response times between depicted rotation in depth and rotation in the picture plane. This difference has not been shown as a simple difference

in the slope of two response time functions (for 'same' pairs), because of another finding. Response times do not increase monotonically with angular difference from 0° to 180° for planar shapes depicted to rotate in depth. Instead response times increase regularly from 0° to 90° of angular difference, and decrease beyond 90°. (A strict difference in slope within a quadrant will be demonstrated in a subsequent experiment: Experiment 3.) Also, it does seem to matter *which* shape is rotated when two different shapes are presented in a pair. This may be taken to indicate that response time differences do not reflect a symmetric process of comparison that is applied to a pair of shapes. Those response time differences could simply be occasioned by foreshortening. Let us address Shepard's objections to the straightforward interpretation of response time differences between depicted rotation in depth and rotation in the picture plane. Planar shapes were used in this experiment throughout, and the conditions of depicted rotation in depth and rotation in the picture plane were identical though run on separate days. Then it cannot be objected that depicted rotation in depth was presented as a simultaneous visual task while rotation in the plane was presented as a task for memory. Nor can it be objected that solid stimuli were used for depicted rotation in depth, while thin plates were used for rotation in the plane. The distinction between the rotation of flat shapes and the rotation of solid shapes in flat depiction may still be an important distinction, but for plane figures at least there does exist a reliable difference in response times between rotation in the picture plane and depicted rotation in depth. The result is not original, but it did require support in face of the objections that have been raised.

3. Experiment 2: Single rotation and tandem rotation in depth

If one accepts the notion that angular difference in depth is confounded with change in discriminability due to pictorial perspective in demonstrations of the mental rotation effect, a question remains: how can one disentangle these supposed influences on response times? A clearer pattern may emerge for response time results when the specific values of geometric variables are taken into consideration. Prospects are better when a measure of change in area is considered: the ratio of the projected area of a shape to its original area. Yet that quantity varies with the cosine of the angular difference of a planar shape from the picture plane. Then changes in response time that are associated with this quantity should be disambiguated from changes due to angular difference. If the two can be disambiguated, changes in response time may finally be associated with this ratio of areas rather than with the magnitude of angular difference. And then changes in response time that have been attributed to angular difference plus 'similarity of shape' or 'complexity of shape' could prove to be interpretable as changes associated with a ratio of areas, or with a ratio of areas plus other basic geometric quantities.

One important distinction between angular difference and the effects of perspective is that perspective effects like foreshortening depend on the profile of a shape *relative to the picture plane*, while the measure of angular difference between two shapes does not depend on the picture plane. Since the angular difference between two planar shapes can be manipulated independent of the angle at which they stand to the picture plane, the

effects of perspective can be disambiguated from those of angular difference. The angular difference between two shapes can be held constant while their angle to the picture plane is varied. There is a brute method and a subtler method to perform this manipulation. The brute method is to set the two shapes in the same plane, and to vary its orientation to the picture plane. In orthogonal projection, the projections of those two shapes will be identical for any orientation of the plane, if the centres of the shapes are positioned at the same distance to the axis of rotation. The subtler method is to separate the two shapes by a small and fixed magnitude of angular difference, and then to vary the yoked orientations of the two shapes with respect to the picture plane. The projections of the two shapes will differ from one another, for most values of their yoked orientation to the picture plane. Let us call this the *tandem rotation* of planar shapes in pictured depth. In this article the orientation of the farthest shape to the picture plane will be used as a measure of tandem rotation. When shapes are subject to single rotation and to tandem rotation at different times, the effects of angular difference can be contrasted with the effects of pictorial perspective. If there is a process of mental rotation based on angular difference, tandem rotation should not produce systematic changes in response times, since in tandem rotation angular difference is constant. But if pictorial perspective is to blame for the changes in response times that are characteristic of the Shepard–Metzler effect in depth, then response times should vary with the magnitude of tandem rotation. The slope of the regression line for the response time function in tandem rotation should be equal to the slope of the regression line for the response time function in single rotation. In this way, a simple manipulation may lead to a better description of the 'mental rotation' of planar shapes in depth.

3.1. Method

3.1.1. Subjects

The eight observers were the same ones who participated in Experiment 1.

3.1.2. Stimuli

Two pairs of flat pentagonal shapes were displayed at different orientations to the picture plane. The first pair consists of Shapes *1* and *3*, (Fig. 4), while the second pair consists of Shapes *2* and *4*. Two shapes from a pair are displayed on each trial; that is, a shape can be matched with a copy of itself or with its partner in a pair. The partners of a pair count as *different*: these shapes are different in projective terms. Each shape is presented in the picture plane, or at one of six orientations to the picture plane: $12.5°$, $25°$, $37.5°$, $50°$, $62.5°$, or $75°$. These can be called pictured rotations in depth, or rotations about the *x*-axis of the picture plane. The pairs of shapes are shown either in single rotation or in tandem rotation. In single rotation, one shape is fixed in the picture plane while the other shape is pictured to have been rotated in depth. In tandem rotation, both shapes are pictured to have been rotated in depth but there is a small initial difference in angle between the two. This may be made clearer by an example. A series of shape pairs (Fig. 5) in single rotation can have the following angular orientations to the picture plane: $(0°, 12.5°)$, $(0°, 25°)$, $(0°, 37.5°)$, $(0°, 50°)$, $(0°, 62.5°)$, and then $(0°, 75)$. A similar progression of shape pairs (Fig. 6) in tandem rotation is: $(0°, 12.5°)$, $(12.5°, 25°)$, $(25°,$
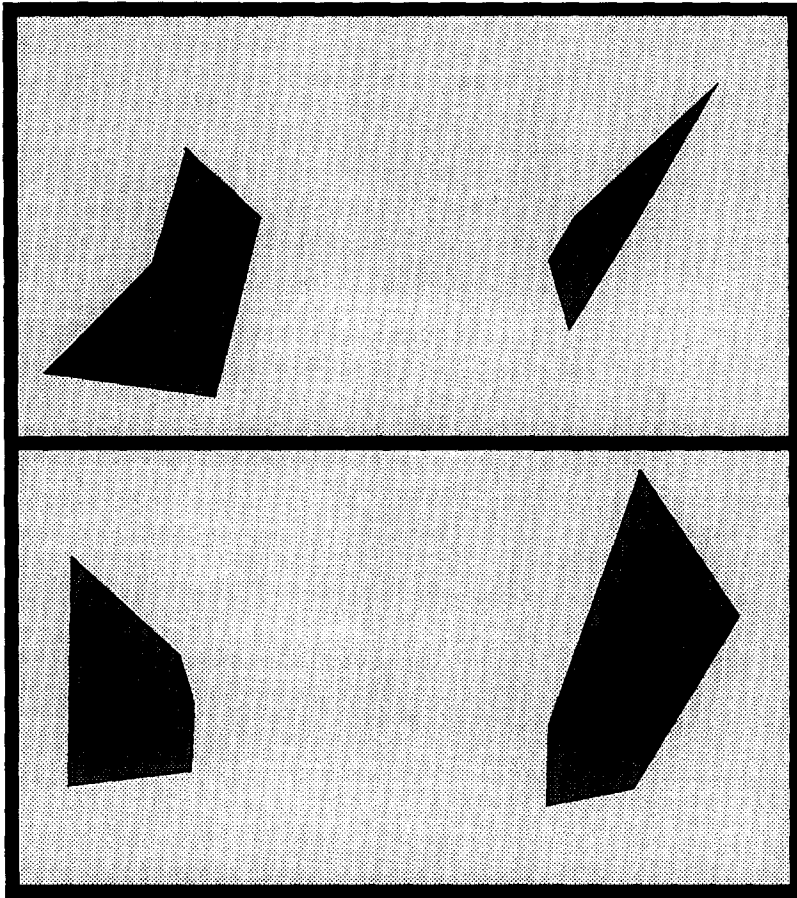
Fig. 4. The stimuli for Experiment 2 consist of four shapes. They are two pairs of shapes that are quite different in projective terms. With reference to Fig. 2, one pair consists of (Shape 1 and Shape 3), while the other pair consists of (Shape 2 and Shape 4).

37.5°), (50°, 62.5°), and (62.5°, 75°). Single rotation alters both the angular difference of the two shapes and the angular difference of one shape from the picture plane. Tandem rotation alters the angular difference of both shapes from the picture plane, but holds constant the angular difference between the shapes themselves.

3.1.3. Procedure

Observers were asked to decide if pairs of shapes were 'the same' or 'different', where 'same' meant unchanging in shape despite rotation or foreshortening. Observers

Fig. 5. The Shepard–Metzler effect in depth has been thought to link increasing angular difference to increasing response times. Here two planar shapes are depicted to have angular differences of 25° (top pair), 50° (middle pair) and 75° (lower pair).

0°   25°

0°   50°

0°   75°

13°   25°

38°   50°

63°   75°
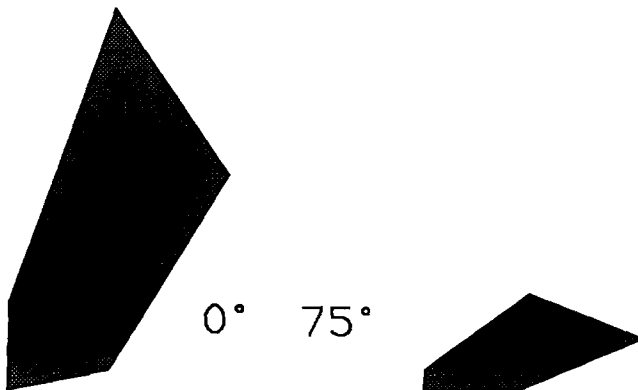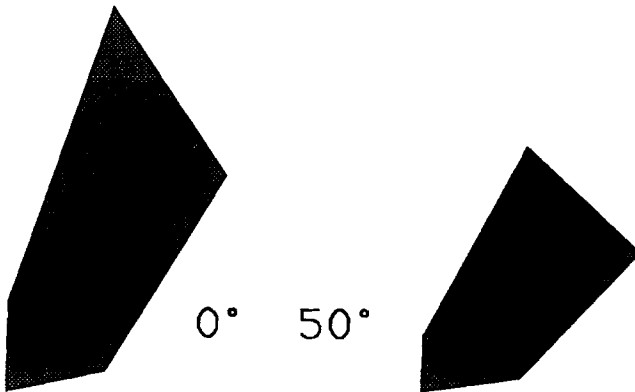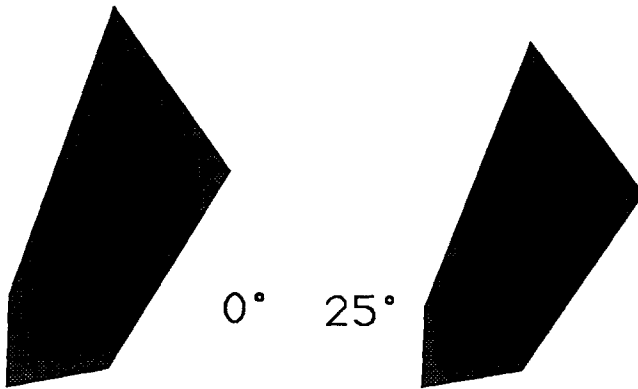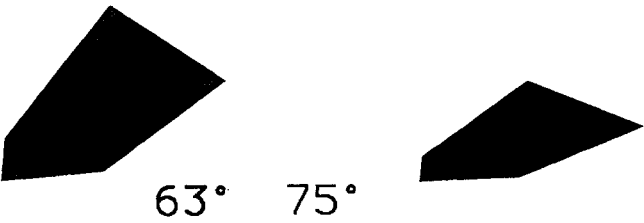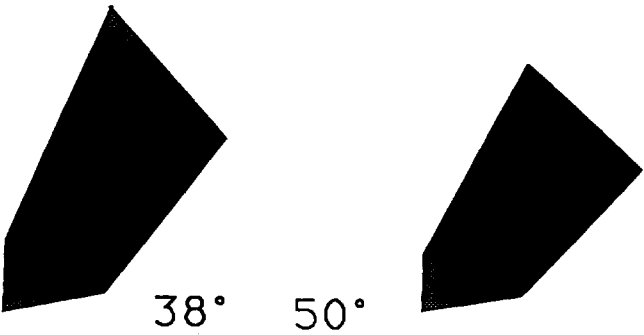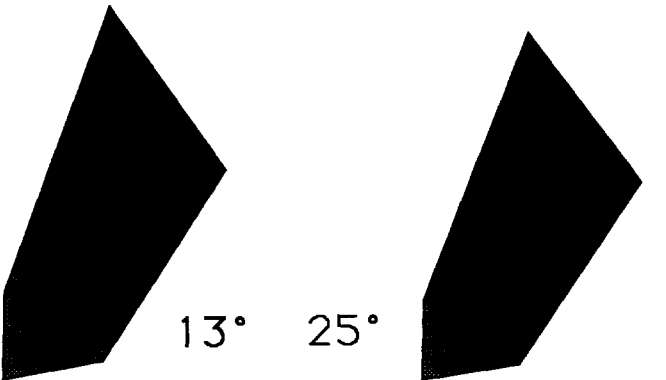
saw 48 distinct pairs of shapes on each day of testing. The order of these pairs was randomized in blocks of 48 trials. Observers saw five blocks of trials each day. The two manners of rotation (single and tandem) were presented on different days. The left–right position of the shape that was rotated farthest from the picture plane was alternated on different days. There are four combinations of these two manners of rotation and two left–right positions. The order of these four conditions was balanced across observers over the four days of testing. There are a total of 960 trials for each of eight observers in the experiment.

The 48 combinations of shapes within a block of trials represent the combination of four factors: *Pair* (two pairs of shapes), *Orientation* (six conditions of pictured rotation), *Same/Different* (the members of the pair count as the same shape or as different shapes) and *Order* (left–right arrangement of each pair of shapes). If the members of a pair are {*a, b*}, then the distinct left–right arrangements of the members in the display are (*a, b*), and (*b, a*). Five *Blocks* of trials were administered each day. Two conditions of left–right precedence of rotation (*Side*) and two manners of rotation (*Rotation Type*) were presented in four days of testing. Then the experiment has a seven-way ($2 \times 6 \times 2 \times 2 \times 5 \times 2 \times 2$, that is, *Pair $\times$ Orientation $\times$ Same/Different $\times$ Order $\times$ Block $\times$ Side $\times$ Rotation Type*) repeated measures design.

3.2. Results

The results for 'same' trials will be reported separately from the results for 'different' trials; the results for 'same' trials will be reported first. A six-way analysis of variance showed a significant effect of change in *Orientation* on the dependent measure of the logarithm of response time ($F(5,35) = 33.20$, $p \leq 0.01$). An effect of *Pairs* of shapes ($F(1,7) = 18.32$, $p \leq 0.01$), and an effect of practice within days (*Block*: $F(4,28) = 16.98$, $p \leq 0.01$) were found to be significant. No other effects were found to be significant. Latencies for Shape *1* were shorter in general than latencies for Shapes *2, 3,* and *4*, though all followed a linear trend with increasing angle of rotation to the picture plane. The mean difference between latencies to Shapes *1* and *3* is 1.25 seconds. It is noteworthy that the conditions of single rotation and tandem rotation (*Rotation Type*) do not have significantly different effects on the dependent measure. Some measures of response time and discriminability are displayed in Table 1 of Appendix C; these vary with angle in both single rotation and tandem rotation.

The slopes of the linear functions that best fit the response time data are similar for conditions of single rotation and tandem rotation (see Fig. 7). The slope of the linear function that relates angle from the picture plane (in degrees) and response time (in

---

Fig. 6. Two planar shapes are separated from one another by a small and constant angular difference; this yoked pair of shapes is then depicted to rotate in depth as a unit. This *tandem rotation* of planar shapes in depth can be used to test the hypothesis that experimentally observed increases in response time are not associated specifically with increases in the angular difference between two shapes, but rather are associated with increases in the angular orientation of those shapes to the picture plane. The fixed angular difference is 12.5° in this example. The closest member of the pair is depicted to have an orientation of 12.5° (top pair), 37.5° (middle pair), and 62.5° (lower pair) to the picture plane.
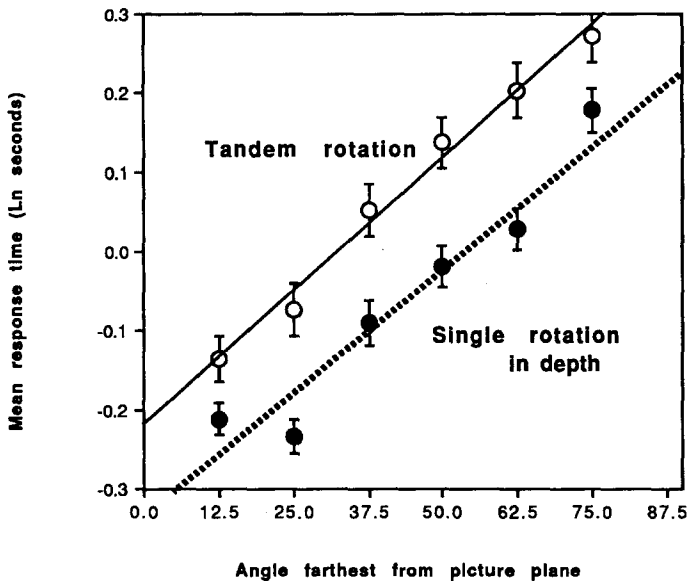
Fig. 7. Mean response time in ln seconds is plotted against the angle of the shape farthest from the picture plane (in degrees), for two conditions of the second experiment. The slope of the regression line for response times to shapes in tandem rotation in depth is very nearly equal to the slope of the regression line for response times to shapes in single rotation in depth. Each point represents the mean of 640 observations, collapsed over the results of eight subjects in the experiment.

ln sec) is $156°/\ln$ sec for single rotation, and $134°/\ln$ sec for tandem rotation. These quantities were computed only for those trials on which a correct response of 'same' was given. (Slopes can also be computed for all trials, both correct and incorrect. These were $156°/\ln$ sec for single rotation, and $148°/\ln$ sec for tandem rotation.) Individual differences in response time will affect the value of such group estimates of slope.

A six-way analysis of variance was also applied to the results for which 'different' was the appropriate response. Again there was a significant effect of change in *Orientation* ($F(5,35) = 14.55$, $p \le 0.01$), an effect of *Pairs* of shapes ($F(1,7) = 40.42$, $p \le 0.01$), and an effect of practice within days (*Block*: $F(4,28) = 8.94$, $p \le 0.05$). There were two significant interactions that involved the factor *Order*: *Pair × Order* ($F(1,7) = 12.53$, $p \le 0.01$) and *Rotation Type × Order* ($F(1,7) = 10.42$, $p \le 0.05$). There was a significant interaction of *Pair × Side* ($F(1,7) = 5.67$, $p \le 0.05$), which was further complicated by a four-way interaction of *Orientation × Rotation Type × Pair × Side* ($F(5,35) = 6.08$, $p \le 0.05$). These last two interactions may be interpreted as the effects of individual shapes under rotation, which are less marked when pairs of shapes are rotated in tandem.

3.3. Discussion

A trigonometric relation shows the difference between an account based on angular difference and one based on increments in area. ''Now the effect on the images

produced by orthographic projection from a plane slanted around the horizontal axis simply amounts to an affine coordinate transformation of the image plane, where the scale of the vertical axis is contracted by a factor equal to the cosine of the angle of slant from the frontal parallel'' (Lappin and Ahlström, 1994, p. 236). When a plane figure at a slant to the picture plane is projected onto the picture plane, the area of the projection can be found as the area of the original, multiplied by the cosine of the angle between the two planes: $\cos \theta = (\text{Area}_p / \text{Area}_o)$. When we consider two plane figures, we may take the difference of these: $(\cos \theta_1 - \cos \theta_2)$. In the limit, the rate of change of $(\cos \theta)$ is $(-\sin \theta)$. By contrast, the rate of change of the linear function that is supposed to exist between response times and angular difference will be a constant value. Suppose that changes in response time were linked to angular differences. Then if the angular difference between two planar shapes were fixed, and their yoked orientation were changed, response times should not change. But suppose that changes in response time were linked to changes in area instead (effectively, to the cosine of the angle of each shape to the picture plane, which represents the degree of compression of the shape). Then if the angular difference between two planar shapes were fixed, and their yoked orientation were changed, response times would be correlated with difference of change in area: they would vary roughly as the relative magnitude of the difference of two cosine functions $(\cos \theta_1 - \cos \theta_2)$. Now suppose that in an experiment (on the Shepard–Metzler effect in depth with planar shapes), the range of angular orientations were both truncated and coarsely sampled, say in 10° intervals from 10° to 80° of tilt to the picture plane (Fig. 8). In the absence of foreknowledge about these functions, $(\cos \theta)$ and $(\cos \theta_1 - \cos \theta_2)$ or $(-\sin \theta)$ would provide an excellent fit to a linear function of angle in the analysis of results. A distinction does emerge clearly in the tandem rotation effect.

Given the number of articles that have been devoted to the Shepard–Metzler effect, and given its status in the literature on cognitive psychology, a simple question comes to mind: why didn't anyone notice this before? The appearance of a significant curvilinearity (say, a quadratic trend) in a polynomial regression of response time over angle might be explained away as the influence of extraneous factors; indeed, such reports have been intermittent and variously interpreted. For example, Koriat and Norman (1985, p. 429) appeal to mental mechanism in explaining the curvilinearity of their response time results for the Shepard–Metzler effect in the picture plane. They propose ''that extensive practice with a visual stimulus results in a memory representation that is broadly tuned, thus enabling efficient stimulus recognition over a relatively wide range of orientations. Small deviations from normal orientation do not require rotation before recognition.'' But still another result very like tandem rotation has been reported in the literature on mental rotation. Massaro (1973, pp. 416–417, esp. Exp. 2) reports experiments in which pairs of circles or ellipses were tilted in depth with respect to an observer. The two shapes were tilted at the same angle to the observer on each trial. Massaro was concerned to show that the perception of shapes in this situation does not involve a process of mental rotation. He found that ''Mean percentage errors and mean RT increased with increases in absolute angle of rotation in a positively accelerated manner...'' (p. 417). He found no difference in his results across two *directions* of tilt in depth away from the picture plane. He continues: ''The results disconfirm the rotation
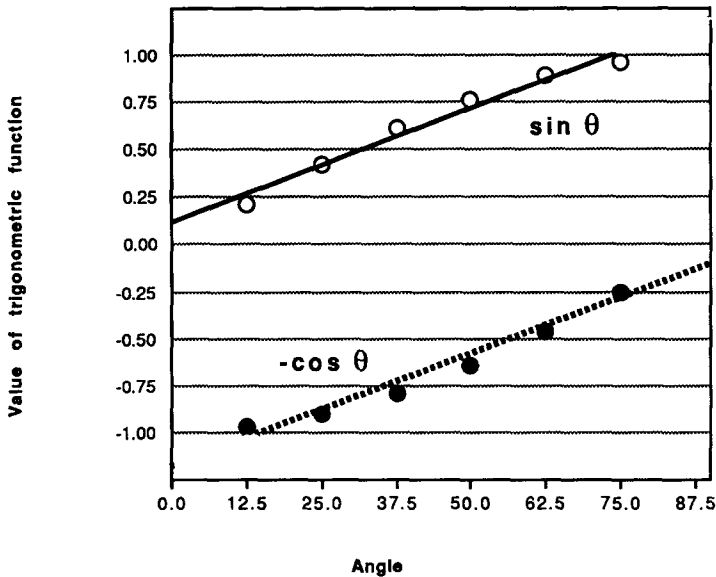
Fig. 8. The data of Experiment 2 can be modelled by simple trigonometric functions. The ratio of the projected area of a shape that is rotated in depth to the area of its unrotated version can be found as the cosine of the angle between the picture plane and the plane of the shape. In tandem rotation, the difference of this ratio between one shape of a pair and its partner follows differences in $\cos \theta$. The rate of change of $\cos \theta$ is $-\sin \theta$ (this is the limit when the difference between angles becomes small; note that signs have been reversed in the graph). When the domain of these trigonometric functions is truncated and coarsely sampled, their plots against θ appear nearly linear in form, with nearly identical slopes. The regression lines have been clipped at ± 1.0 of the ordinates, since that is the range of the functions.

hypothesis as a sufficient condition for the RT functions in this task... the critical variable is degree of rotation from the FPP [fronto-parallel plane], not rotation differences'' (Ibid.). Massaro does not compare observers' results on his task to observers' results in a task more like Shepard and Metzler's. Nor does Massaro claim that a process of mental rotation may be unnecessary to describe the results of Shepard and Metzler's experiments, as well.

Tandem rotation is a simple manipulation which leads to a new description of the Shepard–Metzler effect in depth. Tandem rotation elicits a response time function with a slope not distinguished from the slope of the response time function for single rotation. The new description depends on a variable that can often be measured in the picture plane: a ratio of areas. For single rotation of an identical pair, this quantity is simply the ratio of their areas. This ratio of areas is better expressed as the ratio of the projected area of a shape, to its area before projection. This ratio of areas is also known as the *degree of compression* due to affine transformation in perspective. Note that this quantity is *not* an affine invariant in this situation in one sense: in the sense that it is not an invariant of shape under perspective depiction of rotation in three dimensions. Rather it is an invariant in the picture plane that changes in perspective depiction: hence its

value changes with the severity of the affine transformation that is used to depict rotation in depth.

There may be an effect of perspective on the discriminability of shapes in projection, which increases monotonically with the compression of the shape: an effect of foreshortening. So the crucial ratio of areas varies as the cosine of the angle of a planar shape from the picture plane in single rotation, as a matter of geometry. The linear increase in response times that is characteristic of the mental rotation effect in depth can be described (for the case of planar shapes only) as a linear approximation to a cosine function or a normalized difference of cosine functions. This alleviates a small problem: how response times could appear to differ between single rotation and tandem rotation at the smallest step size of angular difference, where the displays do not differ. (So in this experiment a tandem rotation of 12.5° produces the same configuration as a single rotation of 12.5°. See Figs. 5 and 6.) The answer to this problem is that the linear functions are only approximations to curvilinear trigonometric functions. This interpretation of response times in the mental rotation effect describes the effect of single rotation, and it describes tandem rotation in the same way, by the measure of a ratio of areas. This interpretation will be extended to a half cycle of rotation in the next experiment, where the predictions of an hypothesis of mental rotation and the predictions based on a ratio of areas diverge still further.

4. Experiment 3: A prolonged test with two observers

A systematic effect of tandem rotation on response times has been demonstrated for a group of observers already. One may ask if the same effect of tandem rotation can be demonstrated for an individual, with practice over a long period of time. The tests that will be reported were conducted on two observers over forty days; the range of rotation of the planar shapes was extended to a half cycle (180°). The predictions for response time that are made under the hypothesis of mental rotation diverge markedly from the predictions that are made on the basis of perspective effects, when this second quadrant (90° to 180°) is considered. (In other words, here is a wider look around.)

4.1. Method

4.1.1. Subjects

Two adults from DCIEM were tested. They are: observer KN (the author, 37 years of age, male) and observer VR (21 years of age, female).

4.1.2. Stimuli

The stimuli are two shapes (Shapes *1* and *3* of Fig. 9) and their plane mirror images. These shapes are presented under the conditions of lighting and texture that were described previously. There is one difference: the viewpoint was specified to be further from the picture plane than before (15 units $+z$ instead of 10). The shapes could be rotated either about the z-axis or about the x-axis, that is, rotated either in the picture plane or in depth. The shapes can lie at a variety of orientations from the upright: 15°,
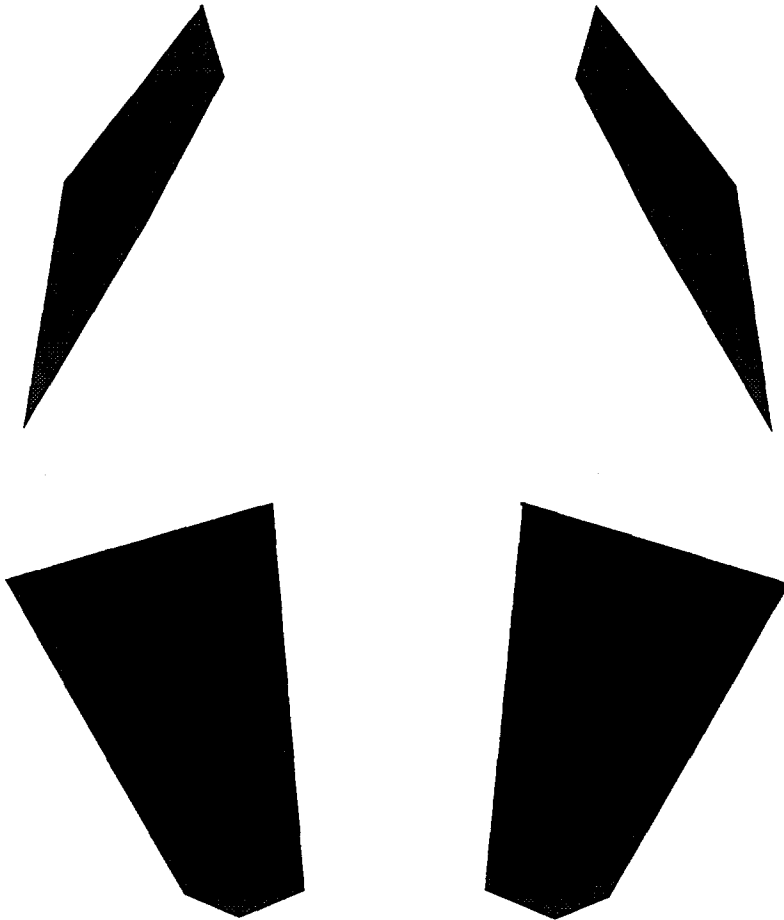
Fig. 9. The stimuli for Experiment 3 consist of four shapes. They are two pairs of mirror-image shapes. Their relative coordinates (in reading order) are:

Shape 1: [(6.38, 8.41), (4.86, 5.24), (2.39, 0.80), (3.20, 6.01), (5.96, 9.93)]
Shape 2: [(3.61, 8.41), (5.13, 5.24), (7.60, 0.80), (6.79, 6.01), (4.03, 9.93)]
Shape 3: [(7.39, 9.11), (8.14, 0.78), (6.79, 0.12), (5.68, 0.63), (2.08, 7.37)]
Shape 4: [(2.60, 9.11), (1.85, 0.78), (3.20, 0.12), (4.31, 0.63), (7.91, 7.37)]

30°, 45°, 60°, or 75° in the first quadrant, or 105°, 120°, 135°, 150°, or 165° in the second quadrant.

Pairs of shapes were presented in two manners: in single rotation or in tandem rotation. Each shape has a fixed orientation to the picture plane. A pair of shapes is in *single* rotation when one shape remains at an upright orientation, and the other takes on successive positions at increasing angular orientation to the first. The angular difference between the two shapes increases. A pair of shapes is in *tandem* rotation when the shapes are separated by a small and fixed angular difference, but they are *yoked* in rotation – they rotate together with respect to the picture plane. The angular difference

between the two shapes remains constant (12.5° here), while the absolute orientation of the shapes changes with respect to the picture plane.

Observers saw 160 distinct pairs of shapes on each day of testing. The order of these pairs was randomized in blocks of 160 trials; observers saw four blocks of trials, for a total of 640 pairs of shapes each day. The *Side* of the shape that was rotated farthest from the picture plane was alternated left to right on different days. The *Axis* of rotation of the shapes (z-axis or x-axis) was changed on different days, as was the type of *Comparison* shape. Either a difference in handedness or a projective difference from the standard shape characterized the comparison shape for a given day. There are eight combinations of the factors of *Side*, *Axis*, and *Comparison*, and these combinations were presented on eight separate days of testing. Eight days comprise one *Series*; each observer completed five series, or forty days in total. The order of days of testing was randomized within these series. Both of the observers endured 25,600 trials.

The 160 combinations of shape pairs within a block of trials represent the combinations of five factors: *Shape* (there are two shapes apart from mirror images), *Quadrant* of rotation (first to 90°, second past 90°), *Orientation* of the shape farthest from the picture plane, *Rotation Type* (single or tandem), *Same/Different* (identical or nonidentical pairs) and *Version* (two arrangements of each pair of a standard and a comparison shape). Four *Blocks* of trials were given each day. The two conditions of left–right precedence of rotation (*Side*), two axes of rotation (*Axis*), and two different classes of comparison shapes (*Comparison*) were presented across eight days of testing. Then the experiment has an eleven-way factorial ($2 \times 2 \times 5 \times 2 \times 2 \times 2 \times 4 \times 2 \times 2 \times 2 \times 5$, or *Shape × Quadrant × Orientation × Rotation Type × Same/Different × Version × Block × Side × Axis × Comparison × Series*) repeated measures design.

4.2. Results

The dependent measure for the first analysis is the logarithm of response time, as before. The application of a single criterion in repeated statistical tests induces an experiment-wise error (of Type I). Experiment-wise error introduces a discrepancy between the nominal level of significance and the real level of significance of a test statistic. A solution is to apply a more stringent level of significance in all tests. In the present experiment $\alpha = 0.001$ will be taken as the minimum acceptable level of significance for rejection of the null hypothesis.

Practice effects are evident in the response times and error counts for each observer. Leone et al. (1993) have reported a gradual diminution of response times with practice in the mental rotation task. Response times decreased regularly in the present experiment, even between thirty-two and forty days of testing (Fig. 10).

An eight-way analysis of variance was applied to each observers' response times; the factor of *Series* was used in the error term; *Side* and *Block* were replicates for these analyses. Some effects were found to be significant in both analyses. There was a significant effect of *Orientation × Axis* (KN: $F(4,16) = 171.79$; VR: $F(4,16) = 90.47$, both $p \leq 0.001$) on the dependent measure of ln response time, as well as an effect of *Rotation Type* (KN: $F(1,4) = 109.85$; VR: $F(1,4) = 123.30$, both $p \leq 0.001$; see Figs. 11 and 12). The *Orientation × Axis* interaction is an expected result: response latencies

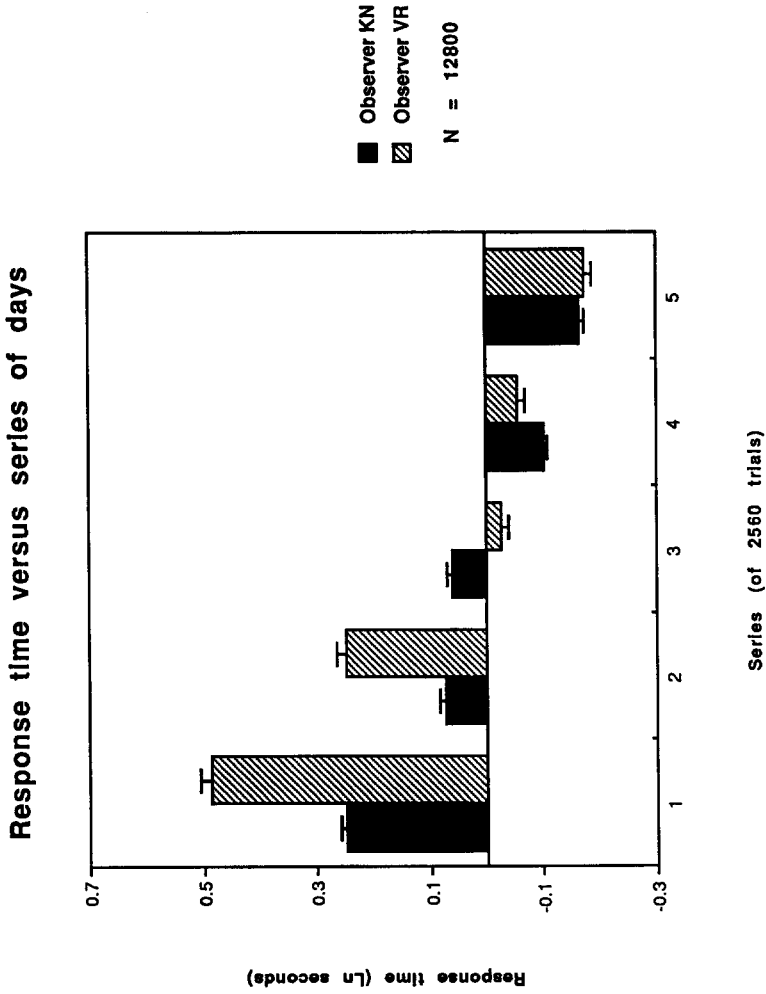Fig. 10. Response time decreases with practice, even after a number of weeks. Mean log response times are plotted for series, that is, for consecutive blocks of 2560 trials in the first quadrant only. Each series represents eight separate days of performance on the task. Only trials for rotations in the first quadrant are included here; the overall trends are not different. Standard error bars are shown.
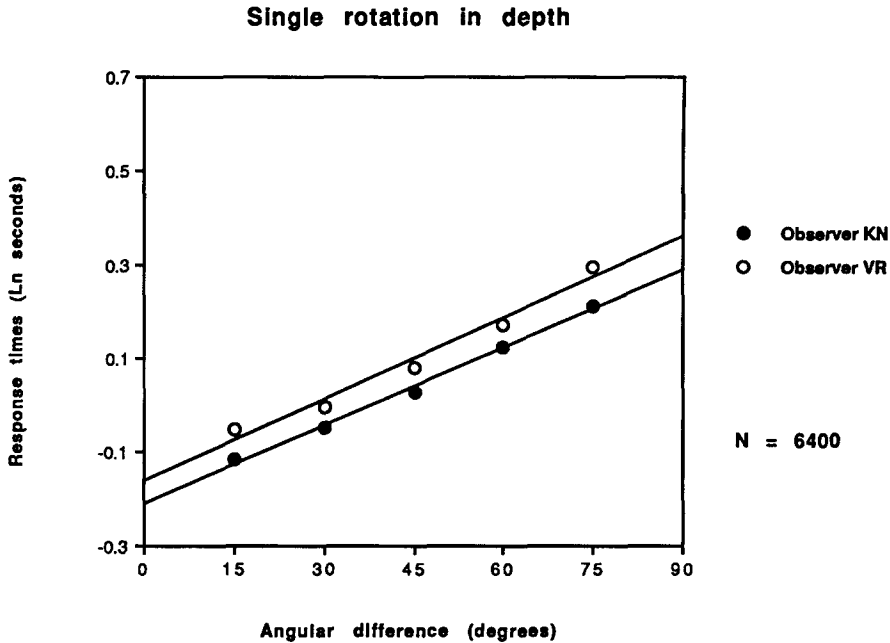
Single rotation in depth

Fig. 11. The Shepard–Metzler effect in depth is illustrated for Experiment 3. Response times increase linearly with angular difference, for both observers. Each dot represents the mean of 1280 observations: these come from a variety of conditions.

in the mental rotation effect are longer for rotations in depth than for rotations in the picture plane (see Fig. 13). This difference occurs when observers are asked to compare shapes that are entirely distinct, just as when observers are asked to distinguish the handedness of shapes. These differences are mirrored in changes of discriminability; the errors that are shown in Table 2 of Appendix D can be used to compute measures of d' for these experimental conditions.

Let us examine the slopes of the linear approximations to the response time functions for single rotation in the picture plane and single rotation depicted in depth, before carrying on with inspection of the ANOVA results. These slopes are estimated within the first quadrant, since it has been indicated for pictured rotation in depth that mean response times need not increase monotonically between the first quadrant and the second quadrant. Only times associated with correct responses were included in the calculations. The slope of the linear function that relates magnitude of depicted rotation in depth (in degrees) and response time (in ln sec) is 141°/ln sec for Observer KN, and 118°/ln sec for Observer VR. The slope of the linear function that relates magnitude of angular difference in the picture plane (in degrees) and response time (in ln sec) is markedly different: it is 294°/ln sec for Observer KN, and 535°/ln sec for Observer VR. (Response times on correct trials can be assumed to be less variable than correct plus incorrect trials, and the variability of times on incorrect trials may be related to the magnitude of times across conditions of angle, so that the better estimate of slope might be obtained from correct trials alone.)
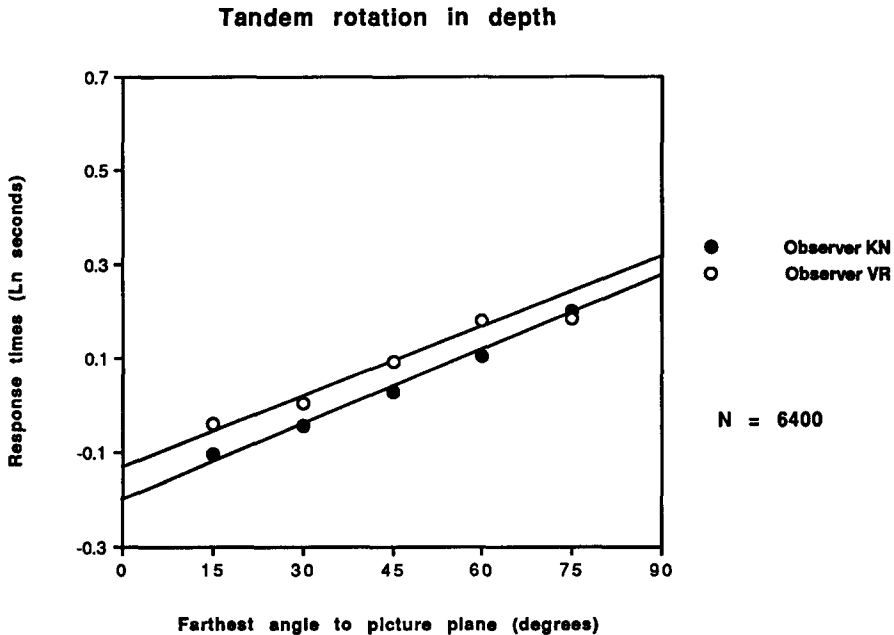
Tandem rotation in depth

Fig. 12. Tandem rotation in depth is illustrated for Experiment 3. Response times increase linearly with angular orientation of the nearest shape to the picture plane, for both observers. Each dot represents the mean of 1280 observations: these come from the same variety of conditions as before. What has been assumed to be the effect of angular difference on response times may be the effect of angular orientation to the picture plane on response times.

Continuing with the ANOVA, response times in tandem rotation were shorter overall than in single rotation. However, this effect was complicated by two interactions: *Quadrant* $\times$ *Rotation Type* (KN: $F(1,4) = 81.07$; VR: $F(1,4) = 327.60$, both $p \leq 0.001$), and *Rotation Type* $\times$ *Axis* $\times$ *Same/Different* (KN: $F(1,4) = 81.82$; VR: $F(1,4) = 111.80$, both $p \leq 0.001$). The significant *Quadrant* $\times$ *Rotation Type* interaction indicates that although response times in tandem rotation were higher than response times in single rotation from 0° to 90°, yet response times in single rotation were substantially higher than those for tandem rotation from 90° to 180°. The *Rotation Type* $\times$ *Axis* $\times$ *Same/Different* interaction is expected on the assumption that there is a tandem rotation effect in depth which does not obtain for rotations in the picture plane.

Responses are relatively fast for rotation in the picture plane, compared to rotation in depth. Among these, the extent and direction of the difference between response times to single and tandem rotation is altered between judgments of 'same' and judgments of 'different'. There is relatively little difference between response times for 'same' and 'different' judgments in single rotation. Responses of 'same' come more quickly for tandem rotation in the plane than responses of 'different', while responses of 'different' come more quickly for tandem rotation in depth than responses of 'same': a comparison across the vertical axis is made more easily for nonidentical pairs in this condition.

Some other effects were significant in the analysis of one observer's results, but not
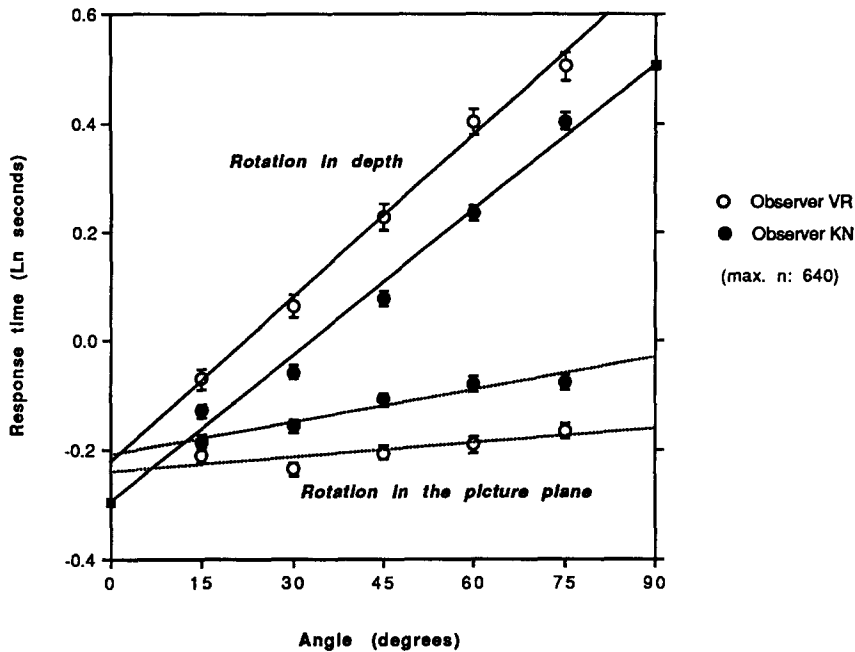
Fig. 13. The slope of the response time function is different for pairs of shapes that are rotated in depth than it is for pairs of shapes that are rotated in the picture plane. Angular difference in degrees is plotted against response time in ln seconds for the results of Experiment 3. Only trials in which 'same' was the correct response are included here. Each dot represents the mean of 640 trials, minus those trials on which an incorrect response of 'different' was made. Standard error bars are shown. For both observers the slope of the response time function is markedly larger for shape pairs that have been rotated in depth.

for the other observer. Some of these were interactions over the *Same/Different* factor for observer KN (*Quadrant* × *Rotation Type* × *Same/Different*: $F(1,4) = 116.07$; *Quadrant* × *Axis* × *Same/Different*: $F(1,4) = 995.40$; *Quadrant* × *Same/Different*: $F(1,4) = 228.55$; *Shape* × *Axis* × *Same/Different*: $F(1,4) = 175.34$; and *Axis* × *Same/Different*: $F(1,4) = 449.21$, all $p \le 0.001$), while a couple of these were simple effects (*Orientation*: $F(4,16) = 163.28$; *Axis*: $F(1,4) = 277.39$, both $p \le 0.001$). The idiosyncratic effects for observer VR were for the factors of *Quadrant* and *Shape*, and the three-way interaction of *Shape*, *Version*, and *Same/Different* (*Quadrant*: $F(1,4) = 124.51$; *Shape*: $F(1,4) = 141.14$; *Shape* × *Version* × *Same/Different*: $F(1,4) = 92.72$, all $p \le 0.001$).

5. Discussion: A trigonometric account

A few simple facts about planar figures that are pictured to rotate in depth can be used to predict patterns of response times in the Shepard–Metzler effect. The basis for these predictions is the trigonometric relation cited previously: when a plane figure at a slant to the picture plane is projected onto the picture plane, the area of the projection
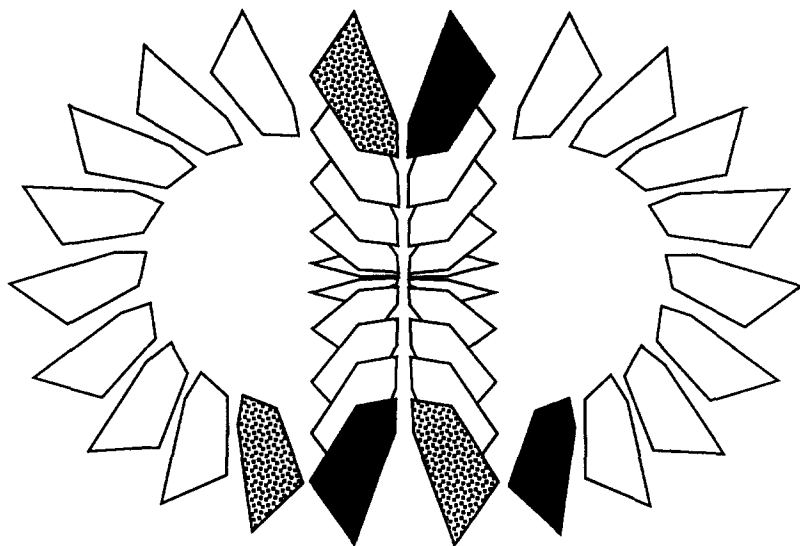
Fig. 14. Rotation in depth has a different consequence than rotation in the picture plane. The shapes at the top of the figure are mirror images of one another, that is, they are of different handedness. A rotation through 180° in depth changes the handedness of both shapes. A rotation through 180° in the picture plane does not change the handedness of either shape. Rotations in these two directions result in shapes of different handedness. The three black shapes are all of one handedness in the picture plane; they are opposite in handedness to the textured shapes.

can be found as the area of the original figure, multiplied by the cosine of the angle between the two planes. Several other facts must be kept in mind. One is that single rotation in the plane, single rotation in depth, and tandem rotation in depth all begin with the same arrangement of shapes. Another is that a single rotation of a half cycle (180°) in the picture plane produces the same arrangement of shapes as does a single rotation of a half cycle (180°) in depth, except for a reflection (see Fig. 14). The third fact is that the arrangement of two planar shapes at 0° is different than their arrangement at 180°. The fourth fact is that when a shape is projected edge-on at 90° to the picture plane, the area of the projection will be at a minimum (i.e., zero). The last fact required is that the area of projection will increase as the shape turns away from that 90° orientation, and that area will increase with increasing angle from that orientation, regardless of the direction in which the shape turns (with respect to the picture plane).

To each of these five geometric facts corresponds an assumption about response times. The first is that one and the same pair of shapes will correspond to a single response time. The second is that the response time to compare two shapes rotated by a half cycle in the picture plane will be the same as the response time to compare two shapes rotated by a half cycle in depth, minus the response time needed to compare two mirror-image shapes. The third is that the response time to compare two planar shapes at 0° is less than the time to compare them when one is rotated to 180°, that is, less by at least the amount of time needed to compare two mirror-image shapes (this is a reflection along a different axis than the previous reflection). Studies of the Shepard–Metzler

effect in the plane have shown that a measurable response time is associated with the comparison of mirror-image shapes, over and above the response time associated with the comparison of matching shapes. This increment in response time (though it may vary somewhat for different axes of rotation) is small relative to the time required to compare two shapes that differ by 90° in the plane. Fourth, because the projections of shapes increase in area as they depart from 90° of rotation to the picture plane (by rotation in depth), response times will decrease as a shape departs from that orientation. Fifth, this decrease in response time will be symmetric over angular orientation in depth about the 90° mark. As they stand these assumptions are not consistent with the response time results of the Shepard–Metzler effect in the plane, though each assumption may seem reasonable on its own.

The problem is that the predicted response times do not add up, nor do they come close. This is apparent even in the assumptions just made, because they incorporate a finding about response times to mirror-image pairs. The response time to compare mirror-image pairs is positive, that is, it is greater than the response time to compare identical pairs of shapes. At least this much must be added to the response time for an identical pair to obtain the response time for an angular difference of 180° in the plane. (Though we may be used to thinking of this transformation as a rotation of 180°, it is also a simple reflection.) And a positive response time to compare mirror-image shapes must again be added to that response time (for an angular difference of 180° in the plane) to equal the response time for an angular difference of 180° in depth. But the overall response time function from 0° to 180° in depth ought to be decreasing and symmetric about 90°, which leaves no room for a positive contribution of these mirror-image comparisons to the total response time at 180° in depth.

This problem gets worse before it gets better. Our third assumption about response times was that the response time to compare two planar shapes at 0° is less than the time to compare them at 180°, by at least the time needed to compare any two mirror-image shapes. A basic finding about the Shepard–Metzler effect in the plane is that response times increase monotonically (or linearly) with angular differences from 0° to 180° (and the increase is substantially greater than the time required to perform mirror comparisons about a vertical axis). Then the response time to compare two planar shapes at 0° is less than the time to compare them at 180° of difference in depth, by a time required to compare two mirror-image shapes, *plus* the time required to compare a 180° difference in the plane. (Again, for a planar shape, a rotation of 180° in depth followed by a reflection equals a rotation of 180° in the plane.) The overall response time function cannot be decreasing and symmetric about 90° for rotations in depth from 0° to 180°, if response times for rotations in the picture plane increase at least monotonically with angular difference from 0° to 180°, and response times are meant to be grossly additive. In other words, our predictions about response times could even violate the triangle inequality when paired with basic claims about the mental rotation effect in the plane. Then let us re-examine our assumptions.

Three of the assumptions depend on the identity of picture pairs. These assumptions seem uncontroversial, as does the fourth, which establishes a maximum response time for an orientation that is difficult to discern. The fifth and last assumption is that the increase in response times from 0° to 90° should be symmetric to the decrease in

response times from 90° to 180°. Here 'symmetric' can mean two things. It can mean that the response time function has the same form across each domain (eg., linear with a fixed slope), and it can mean that the heights of the two functions are identical at corresponding points (in contrast to the case where the intercepts of the linear functions are different, though the functions have the same slope). A change in the form of the function would have a strong interpretation, while a change in the height of the function would not. Then we may postulate that the height of the response time function from 90° to 180° in depth is augmented by the response time attributable to the mental rotation effect through 180° in the plane, plus the response time due to reflection about a particular axis. The response time function will no longer be symmetric to that from 0° to 90°, but the two functions will be the same in form.

Our assumptions have not yet specified the form of the response time function. My hypothesis is that for pictured rotation in depth of a single member of a pair of planar shapes, observers' response times will vary with the compression in area induced by the depiction of rotation (in the long run). The response times will vary with the ratio of the projected area to the original area, that is, with the cosine of the angle depicted between the two shapes. Response times will vary *inversely* with the cosine of angular difference from 0° to 90°; it takes more time to judge a shape that is tilted to a large angle from the picture plane than it takes to judge a shape that is tilted only a little from the picture plane. Response times for rotation from 90° to 180° in depth will follow the same function with a different intercept and a different sign: they will follow $(\cos \theta) + C$. In advance of information about response times for reflection and for mental rotation of a half cycle in the plane, the constant C may be set to one (i.e., the cosine of 180 degrees). It is conceivable that the effect of a mirror reversal of one figure in a rotation through the second quadrant would be some multiplicative effect rather than an additive one, so that the proper function would be $C(\cos \theta)$ or some other form than $(\cos \theta) + C$. This seems unlikely at first, and would need further justification. An overview of these predictions is presented in Fig. 15.

There are marked differences between these new predictions and the predictions that are usually made for the Shepard–Metzler effect in depth, on the hypothesis of a mental process of rotation. That hypothesis concerns angular differences, and does not predict changes in response time when angular difference is held constant, as in the tandem rotation effect. Under the hypothesis, the maximum of the response time function occurs at 180° of difference, and response times decrease symmetrically and linearly from that mark. The existence of a maximum response time at 90° is not predicted for planar shapes; neither is an asymmetry about the maximum for the total response time to a pair of shapes. Then compare the new predictions illustrated in Fig. 15 with the response times plotted in Fig. 16 (also see Table 2 of Appendix D).

My claim so far is as follows: a piecewise continuous function, based on the cosine function, describes response time differences when observers compare flat shapes (where one of these shapes is pictured as rotated in depth). On first thought, this claim contrasts starkly with Shepard and Metzler's (1971) claims about response time differences when observers compare complex shapes in perspective (where pairs of solid shapes are pictured as rotated in depth). Shepard and Metzler describe the angular difference function of response time as a linear function, one sometimes increasing and
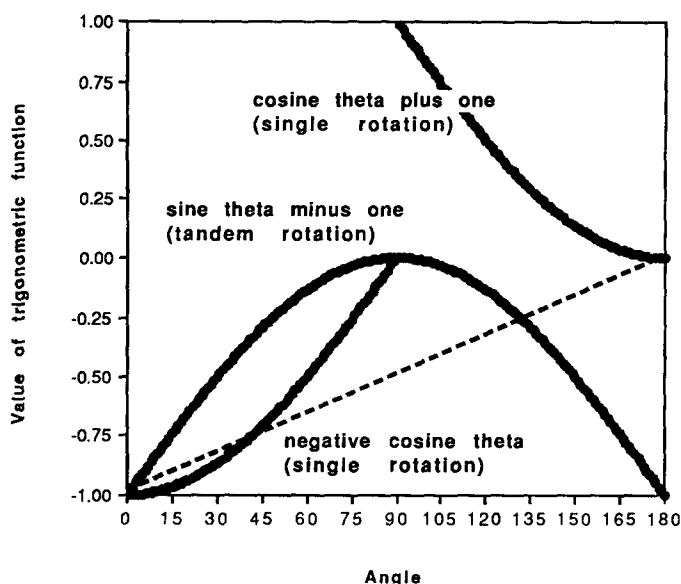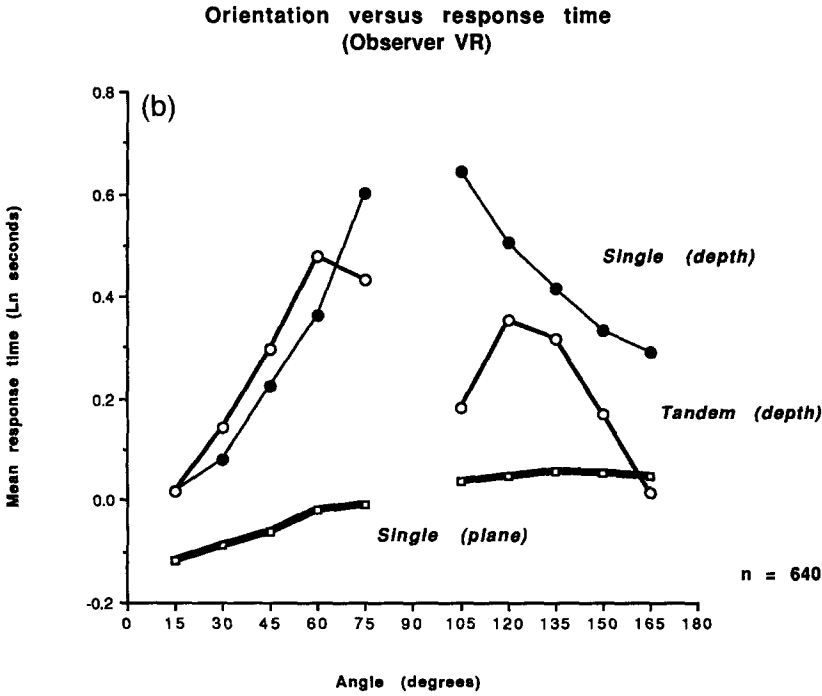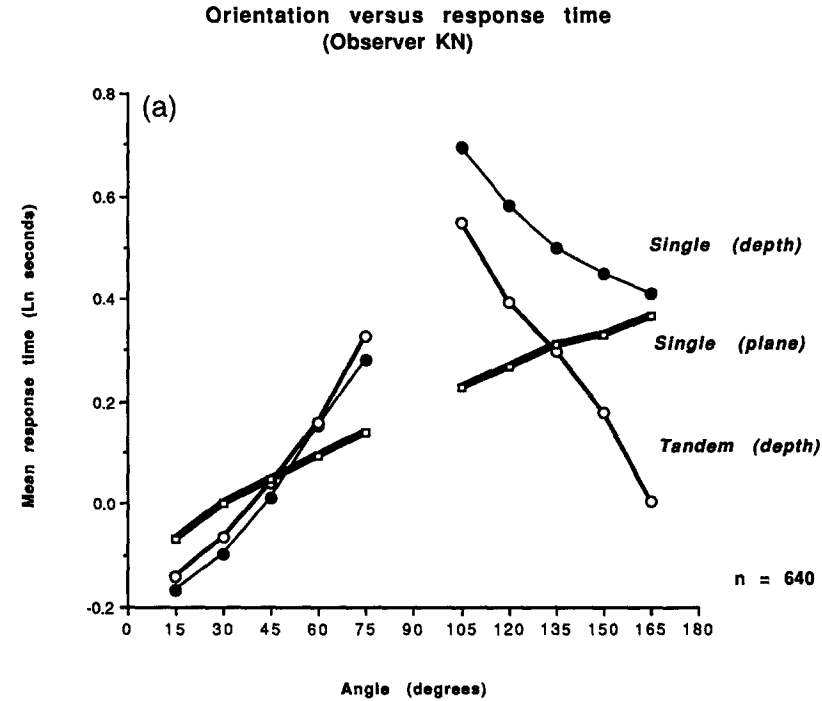
Fig. 15. Response time predictions for the comparison of planar shapes in depth can be drawn from the supposition that those times vary inversely with the ratio of the area a shape at a slant projects, to its original area. This quantity varies with the cosine of the angle θ between the picture plane and the plane of the shape. In the first quadrant (0° to 90°), response times for single rotation in depth should vary as negative cosine θ. In the second quadrant (90° to 180°), response times for single rotation in depth should vary as cosine θ plus an arbitrary constant, which has a maximum value of one. Response times for tandem rotation in depth are adjusted by a constant so that predictions coincide at 0° and 90°. The response time predictions for single rotation in depth include an arbitrary constant. The value of this constant is given by the response time allotted to the 'mental rotation' of a shape in the picture plane (shown as a dashed line) plus a response time to compare mirror-image shapes.

sometimes decreasing in slope. In other words, they describe a piecewise linear function: a triangle function. They are emphatic that the branches of this function are precisely linear: they find no significant quadratic or cubic trends of response time over angle for the results of their extensive, well-conducted experiment. Assume that Shepard and Metzler are right: observer's response times are well described as a piecewise linear function of angle, at least for the comparison of solid shapes in perspective. In particular, the function does not consist of a straight-line approximation to a cosine function.

Then what could be the relation between response times to flat shapes pictured in depth, and response times to solid shapes in perspective? The response time functions for complex solid shapes in perspective can be considered a *consequence* of the response time function for flat shapes in perspective. That is, the response time function for complex solids in perspective can be thought to be *composed* of or built from many response time functions for flat shapes in perspective.

Any solid shape can be thought to be composed of many plane sections or facets. Each of these plane facets makes a dihedral angle with the picture plane; the magnitude

Orientation   versus   response   time
(Observer KN)

Angle   (degrees)

Orientation   versus   response   time
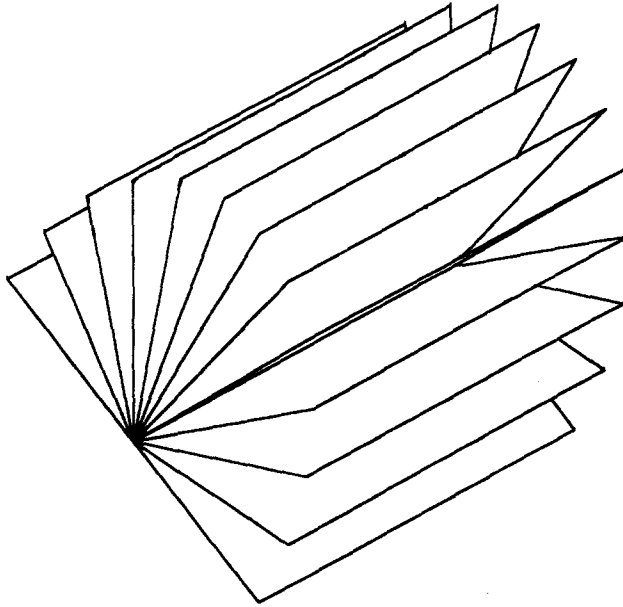**(Observer VR)**

Angle   (degrees)

Fig. 17. The normals to a complex solid shape can be considered to specify a distribution of orientations with respect to the picture plane. The orientations to the picture plane of the facets of a complex solid shape will range over a half cycle, that is, from 0° to 180°. This suggestion of a variety of planes is offered as an aid to the imagination.

of this angle is the orientation of the facet to the picture plane. Any solid shape can be considered for the collection of orientations its facets make to the picture plane (Fig. 17). That is, for this illustration we require no other information about a complex solid shape than the orientations of its parts to the picture plane. Such orientations fall between 0° and 180° from the picture plane. Once the orientations of these plane facets to the picture plane are known, a response time function is associated with each plane facet, as in the account given earlier. Each of these response time functions is a piecewise continuous cosine function (Fig. 18. This piecewise continuous function is the simple extension of the function associated with single rotation that is displayed in Fig. 15). Plane facets that differ in orientation to the picture plane are assigned response time functions (i.e., angular functions of response time) that differ in *phase*. Plane facets that differ by 30° in orientation are associated with angular functions of response time that differ by 30° in phase (as in Fig. 19). Then the response time function for a complex

Fig. 16. Mean response times are plotted against angular orientation from 0° to 180° for the results of Experiment three. Each dot represents the mean of 640 observations. Conditions of single rotation in the picture plane are marked by a thick line and open squares; conditions of single rotation in depth are marked by filled dots; and conditions of tandem rotation in depth are marked by open dots. Data for each of the two observers are plotted separately. It should be noted that the 105° condition for tandem rotation may be unrepresentative.
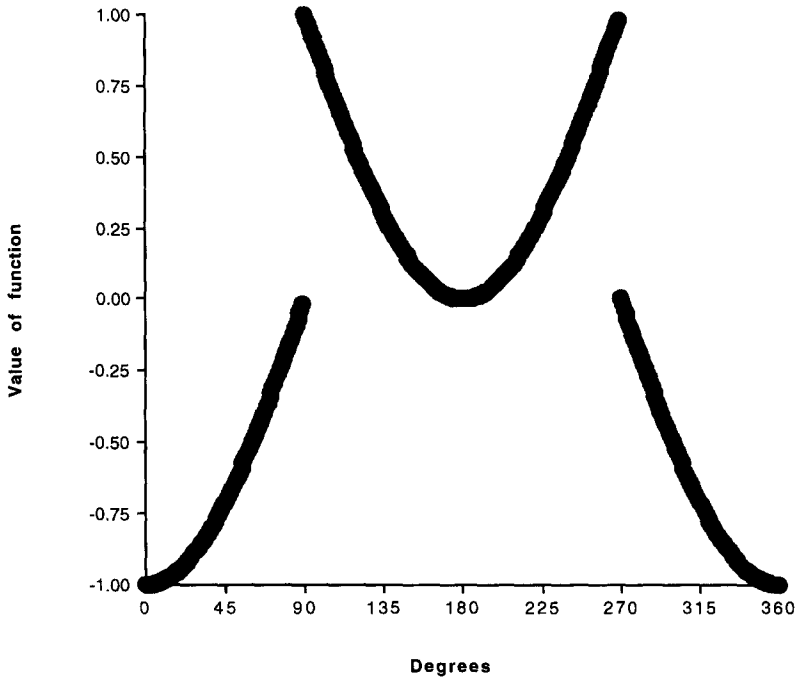
Fig. 18. The piecewise continuous function displayed here is the continuation of the piecewise continuous function that is displayed in Fig. 15. It is proposed that comparison response times will provide a model for this function, specifically where planar shapes are depicted in single rotation in depth.

solid shape can be predicted. The form of that function is the normalized discrete sum of the response time functions for the planar facets or parts of the shape.

This idea can be expressed more generally. Consider the normals to a very complex solid shape, i.e., the directions perpendicular to tangent planes of its surface. The lines of these directions specify a distribution of orientations to the picture plane. Such orientations range from 0° to 180°. A flat shape specifies one of these orientations: what if indefinitely many different orientations are specified? The angular function of response time associated with a complex solid shape is the integral with change in phase over the interval from 0° to 180°, of another angular function of response time: the piecewise continuous function that is associated with a flat shape (Fig. 20). The function associated with one orientation (i.e., with a flat shape) is a piecewise continuous function based on the cosine function. The integral of this function with change in phase from 0° to 180° is a piecewise continuous linear function. (The integral is evaluated in Appendix E.) It is a triangle function, a fraction of the amplitude of the original function (Fig. 20). By integration, Shepard and Metzler's response time function can be obtained as a generalization or consequence of the response time function set out here for flat shapes. Recall that this response time function for flat shapes is based on the degree of compression, or the change in area of a flat shape as it is depicted to rotate in depth. So Shepard and Metzler's response time function can be derived from the degree of

Degrees
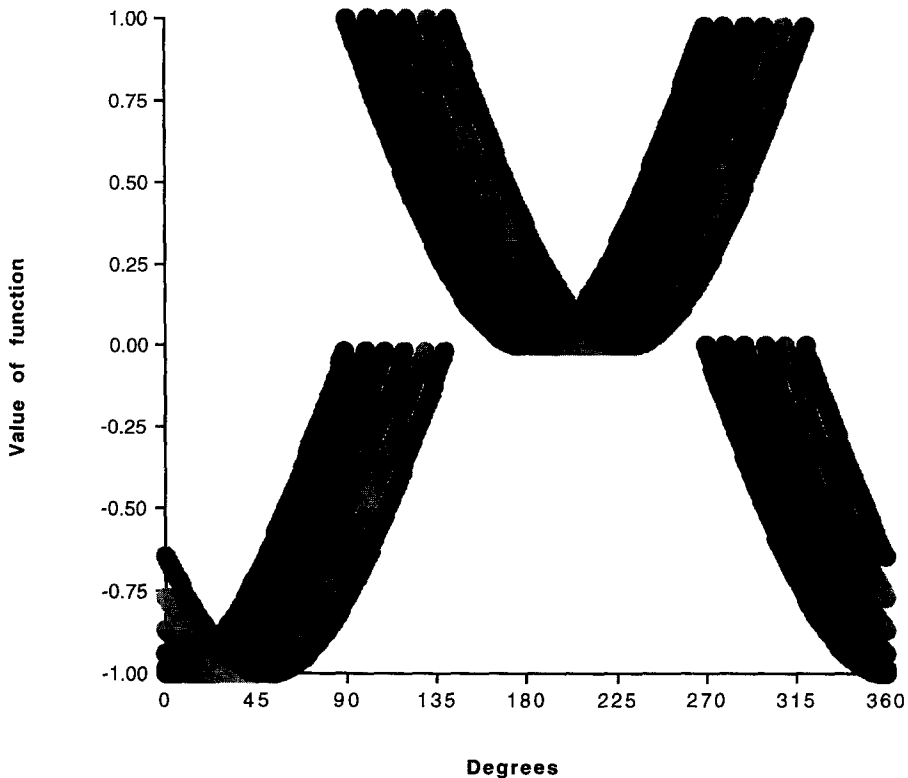
Fig. 19. The facets of a complex shape may differ in orientation with respect to the picture plane. It is proposed that a piecewise trigonometric function predicts response time differences for each of these plane facets. Differences in the orientation of the facets correspond to differences in phase in the associated trigonometric function of angle. Several trigonometric functions that differ in phase over 45° are shown; they are marked in shades of gray. Colloquially, differences in orientations among the parts of a complex shape 'blur' the proposed response time function over 180° of angle.

compression, that is, from the severity of affine transformation, that many parts of complex shapes undergo in a display. (As outlined earlier, some elementary geometric assumptions must be made as well.) This may be a step to explanation of the piecewise linear form of the Shepard–Metzler function of response times (which is a puzzle for Cohen and Kubovy, 1993, p. 381, as well as for many others). Two questions must be addressed before this idea can be extended to particular shapes of intermediate complexity: (1) do occluded surfaces count?, and (2) do many surfaces of the same orientation count for more than a single surface?

The results of this experiment reinforce and extend the results of the preceding experiment. Response times in the Shepard–Metzler effect in depth may be better predicted by ratios of areas than by magnitudes of angular difference. The differences between the two sets of predictions can be made clear by experiment: tandem rotation, and the single rotation of planar shapes in depth from 90° through to 180°, are two
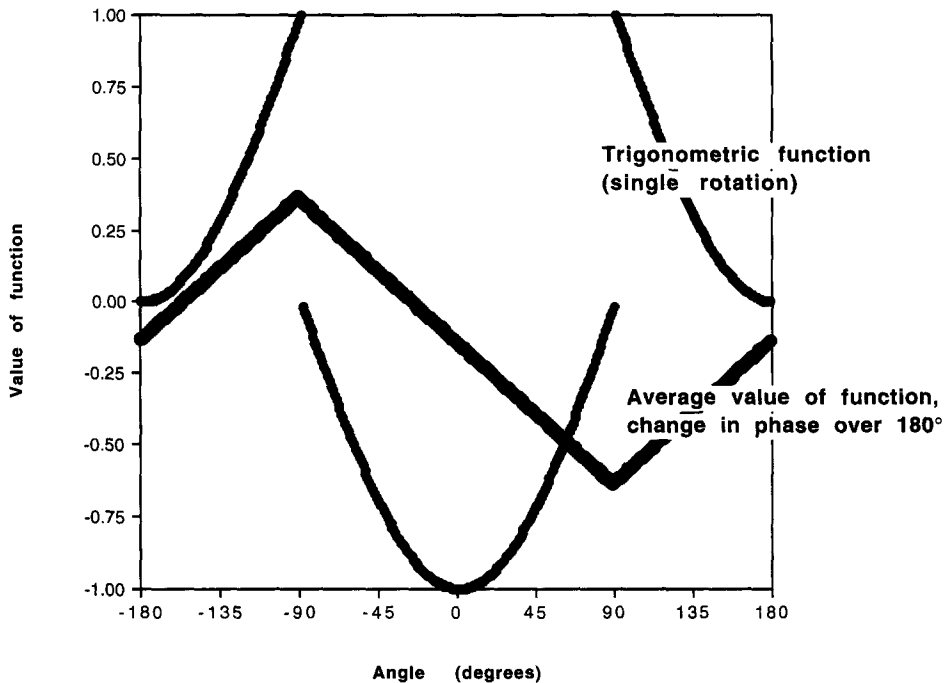
Fig. 20. The Shepard–Metzler response time function for complex shapes in depth is a definite integral of the proposed response time function for single rotation in depth, with change in phase over a half cycle (180°). The characteristic piecewise linear form of the Shepard–Metzler function can be considered a consequence of trigonometry. It is a consequence of the compression in area of the parts of the complex shape, which compression is due to the depiction of the parts in perspective. Only one substantive psychological assumption is involved: that response time differences are proportional to the magnitude of this compression.

examples of manipulations for which the two sets of predictions diverge. They mark the Shepard–Metzler effect in depth as distinct from the Shepard–Metzler effect in the picture plane. They may not imply that the two functions are necessarily different in general form: response times for rotations in the picture plane could follow a trigonometric function of angle as well (this may depend on the shapes in question), or else they could follow a linear function of angular difference. In either case the two conditions of depicted rotation in depth and rotation in the picture plane will be different for planar shapes, and it is the effect of this distinction that we set out to show at the beginning. The perception of plane figures set in the picture plane is not to be explained in the same way as the perception of shapes depicted at a slant, or the perception of shapes susceptible to foreshortening. In the Shepard–Metzler effect in depth, one may measure a simple geometric quantity in the plane of the display, and find that observers' response times vary with the magnitude of that quantity. No kinematics of the mind is required to explain these variations in response time. An hypothesis of mental rotation is neither necessary nor sufficient to explain response times for the simultaneous comparison of planar shapes depicted in depth.

Still one may be left with some question whether to consider the relation of response

time to the ratio of areas to be mediated by the direct apprehension of a ratio of areas, or to consider the relation to be mediated by a process of the discrimination of compression. That question presupposes a psychological mechanism, or a psychological process on the model of physical processes, and it was that confusion of a psychological process with physical processes which spawned the hypothesis of mental rotation. "In general it seems to me better not to try to create a model ..., but to leave these facts as they were born, in an inventory of invariances under various reductions" (Boring, 1952, p. 147). I suggest that additional experimentation might be focussed on a detailed description of variations in response time with the magnitude of geometric invariants and geometric transformations, rather than in search of an explanation based on current notions of psychological mechanism. "Since everything lies open to view there is nothing to explain. For what is hidden, for example, is of no interest to us." (Wittgenstein, 1953, p. 50, para. 126.)

6. General discussion

"The blessed will not care what angle they are regarded from, having nothing to hide."
  (Auden, 1966, p. 241)

The results of the experiments have addressed several different questions about the perception of planar shapes pictured to rotate in depth. The first experiment compared response times for rotation in the picture plane with response times for pictured rotation in depth. It was expected that a difference would be found in the slope of the linear function of response time over 180° of angular difference. Response time differences were found between the two conditions, but not exactly as expected, since response times did not increase monotonically past 100° of angular difference for pictured rotation in depth. The Shepard–Metzler effect in depth is different for planar shapes than for block shapes in that respect. Again in this experiment, response times depend on the particular shape that is pictured to rotate in depth. It was noted that the measure of change in area should be disambiguated from depicted angular difference, since the measure of change in area varies directly with the cosine of the depicted angle of a shape from the picture plane.

A simple manipulation was used to disambiguate angular difference from change in area for the second experiment: the tandem rotation of planar shapes. A linear fit to a response time function was found for pairs of planar shapes in tandem rotation (in depth) under the same conditions that a linear fit to a response time function was found for pairs of shapes in single rotation. The slopes of the two functions were not found to be different. It was supposed that the measure of change in area (the ratio of projected area to original area) might provide an apt description for these results. Then response times would vary inversely as the cosine of angle from the picture plane under conditions of single rotation, and they would vary inversely as the normalized difference of cosines of two angles from the picture plane (roughly, as the negative sine of their median angle) under conditions of tandem rotation. The linear approximations of these

trigonometric functions would be expected to have the same slope under the conditions of the experiment. However, the main finding of this experiment is that monotonic increases in response time can be obtained in the Shepard–Metzler effect, when depicted angular differences are held constant.

A long series of trials was conducted on two observers in the third experiment. A difference in the slope of the response time functions was found between rotation in the picture plane and depicted rotation in depth. A reversal in the relative magnitude of response times was found for single and tandem rotations between the first quadrant of rotation (0° to 90°) and the second quadrant (90° to 180°). Other variations in response times are matched by variations in a measure of the ratio of the projected area of a shape to its original area. That is, the measure can be used, together with some basic facts of geometry, to make coherent predictions of response times in the Shepard–Metzler effect. The effect of these results is to encourage acceptance of a new description for the Shepard–Metzler effect in depth. This description, cast in terms of values of invariants and simple trigonometric functions of response time, provides an alternative to the current interpretation cast in terms of imagined angular differences and linear functions of response time. Yet there is a sense in which the description offered does not supersede the old interpretation. The new description concerns the degree of compression to which a pictured shape may be subject. A measure of this degree of compression does not, and is not meant to provide a criterion of identity for shapes. It does not provide an answer to a basic question, the one that a process of mental rotation was supposed to answer: how do we compare shapes in perspective? A mental process of rotation never supplied a criterion (psychological or not) for the identity of shape, either. Then it remains imaginable that response time differences are epiphenomenal here: both to the comparison of shape by observers, and to hypothetical 'processes' of comparison internal to those observers.

As mentioned in the introduction, Shepard offers some problems or concerns for an account of the Shepard–Metzler effect based on the perception of geometric invariants. Shepard's claims are based on the hypothesis that mental rotation bears a simple relation to response time, as speed of rotation does with respect to time in kinematics. Each of his four claims can now be revised for the case of planar shapes in pictured depth. First, it is not enough to say that response times to decisions of 'same' or 'different' increase linearly on average as the angular difference between two planar shapes increases. Instead response times increase with the cosine of the angular difference between a planar shape and the picture plane, for the Shepard–Metzler effect in depth. Second, response times do not always increase as far as an angular difference of 180°, past which response times decrease at the same rate to 360°. Such changes in response time depend on the geometry of the pictured shape. Response times increase to 90° from the picture plane for the single rotation of a planar shape in depth, after which response times decrease from 90° to 180° (though the function is not continuous). Third, the slope of the function of angular difference versus response time is not identical for rotation in the picture plane and for perspective pictures of single rotation of planar shapes in depth. The two cannot be equal, given some facts of geometry and some simple assumptions about the mental rotation effect. The slope of the function for rotation in the picture plane is less than the slope of the linear approximation to the function for single rotation

in depth, in each quadrant. Fourth, it is misleading to say that the degree of similarity between the perspective traces of the two shapes in the picture plane has no independent effect on response times. The perspective traces of shapes in the picture plane can be used to predict response times, for identical pairs in either single rotation or tandem rotation in depth. But a degree of similarity is not at issue, unless a ratio of areas counts as a measure of the similarity of two perspective traces. A ratio of areas is a definite quantity unlike a degree of similarity, which could denote similarity along any dimension whatsoever. Then the problems or concerns with an approach based on geometric invariants become problems or concerns for an approach based on angular difference. They are:

Result 1.   Response times increase monotonically with the angular difference in orientation of the plane of flat shapes from the picture plane, not with their angular difference from one another.

Result 2.   The rate of this increase in response times is substantially greater for depicted rotations in depth than it is for rotations in the picture plane.

Result 3.   The increase in response times follows a trigonometric function given by simple geometric properties of the shapes as projected onto the picture plane. The range of this increase depends upon the shape in question; for planar shapes response times increase to 90° from the picture plane.

Those are not the only problems that confront an account of the Shepard–Metzler effect based on angular difference. Six other questions suggest themselves:

1. Why do some shapes take longer to rotate than others?
2. How do observers rotate shapes in the right direction – that is, the short way around and not the long way around?
3. Why do responses of 'different' to non-matching (specifically, enantiomorphic) pairs of shapes show the same rate of increase in response time across angular difference, as do responses of 'same' to matching shapes?
4. What accounts for the regular increase of response times associated with the tandem rotation effect?
5. Why are response times to shapes in tandem rotation (up to 90°) longer than response times to shapes in single rotation, in the long run?
6. Why are response times to planar shapes in single rotation not symmetric between quadrants?

This list is not intended to be exclusive or exhaustive. The first question has been addressed in two ways (from the standpoint of kinematics, and from the standpoint of dynamics), neither of which has been wholly satisfactory. Those (like Shepard) who emphasize a kinematic interpretation of the Shepard–Metzler effect, claim that differences among shapes can be explained as differences in the complexity of those shapes or as differences in the familiarity of those shapes. Presumably the more complex a shape is, the more difficult it is to rotate. Yet an independent account of complexity – or

better, a definitive metric for complexity or familiarity – is lacking. The answer to the first question, in terms of an explanation by the discrimination of invariants, is that these shapes have different geometric properties, and different response times reflect the differences in *specific* properties.

The second question becomes important as observers compare pairs of shapes that are quite different in orientation. How should an observer know to rotate a shape through, say 160°, and not 200° in the opposite direction? (The question is raised both by Corballis, 1988, and by Jolicoeur, 1990, p. 399.) Appeal to a principle of least action (Shepard, 1987, p. 87) may be used to justify a correct choice of direction, but this does not constitute an explanation. Even for small angular differences, why should one not start out with the wrong direction of rotation, at least sometimes? Were this to happen, the calculated rate of increase in response times with angular difference would be dampened. Indeed, observers can be instructed to rotate shapes 'the long way around' in the picture plane at least, as Metzler and Shepard (in Shepard and Cooper, 1982, p. 53) have shown. This question does not arise for the alternate explanation, which does not involve angular difference.

The third question concerns response times to different pairs of shapes. It is clear how mental rotation might have been thought to aid the comparison of two shapes in space: one test of congruence is to bring two shapes into superposition by rotation or translation. This procedure would ensure that the initial difference between the shapes is not the result of foreshortening, or other effects of perspective. Yet is this procedure necessary for a pair of shapes whose difference is apparent at first glance – for example, where one shape is almost round and another is long and in places concave? A process of mental rotation does not seem to be necessary for the naming and categorization of distinctive letters of the alphabet when they are rotated in the picture plane, at least (White, 1980). Still, this does not explain why such a pattern of response times does occur – or why a process of mental rotation should be thought necessary – for shapes that are clearly distinct, but which differ in pictured orientation by a rotation in depth. Part of the answer may lie in the assumption that such a process is necessary to *compare* shapes. The pattern of response times could be associated with the discernment of shapes, rather than their comparison. Under an account based on invariants, the difference in response times between 'same' judgments and 'different' judgments depends on the geometric criterion that distinguishes what counts as the same from what counts as different. The remaining questions arise from the results of the present studies; these results were not predicted by an account based on angular difference.

The general finding of these experiments concerning the Shepard–Metzler effect in depth is that response times are not associated with the magnitudes of angular differences that are depicted, but they are associated with changes in invariants that can be measured in the picture plane (i.e., with a measure of the degree of compression due to perspective). Can the findings of these experiments, in which flat polygons have been presented, be generalized to experiments with other shapes, say Attneave's complex polygons (Attneave and Arnoult, 1956), Shepard's block assemblies (R. Shepard and J. Metzler, 1971), or Rock's twisted wire frames (Rock et al., 1989)? Farah et al. (1994, p. 340) sought to address the effect of such stimulus differences on the recognition of shape, but did not stop to measure differences in shape. Angular differences due to

rotation are easy to measure and depict for these shapes, but how can changes in invariants be measured and compared? *Various* techniques of measure are appropriate to different situations; the technique of measuring the area of a square is unlike the technique of measuring the volume of a mountain. The original conception of mental rotation abstracts from particular shape: the rotation of a hand can be compared immediately with the rotation of a letter of the alphabet. What this shows is that the original conception of mental rotation offers no account of form perception at all, since in our environment rotation is neutral to shape. The point can be expressed in another way: explanation of the Shepard–Metzler effect in depth has concentrated on operations or transformations, to the neglect of invariants of shape under those transformations. Geometry provides techniques for the assessment of those invariants. Geometry provides *families* of techniques for the assessment of invariants in vision research. Then the task of generalizing the present findings across different classes of shapes is a geometric task first of all. And while (or when) the results of different experiments with various stimuli can be compared by the astute application of geometry, one should not expect to measure one invariant by one method across many different situations. Nor should one expect a technology to double as a complete psychology of vision, or as an epistemology.

The distinction between psychology and geometry has been effaced too long in the theory of vision. One reason is that investigators may fail to recognize that some criteria for psychological phenomena are either objective or conventional: there are criteria which are *objective* in that they specify states not inside the head or brain or mind, and there are criteria which are *conventional* in that they are not set by any one person or ego or observer. Given conventions about perspective, the criterion for two profiles to represent the same shape at a slant in a picture is an objective criterion. It can be decided with a straight edge, or with a ruler and a calculator, that two shapes count as the same in perspective. Then either observers adhere to these criteria in pronouncing 'same' and 'different', or they do not. The criterion for two profiles to represent the same shape is not, and never has been, a matter of comparing the shapes of mental images. "Although behaviour manifesting understanding is not itself understanding, but evidence for it, what this is evidence for is an ability, not a state, i.e. not a persisting mental structure in an ethereal medium. Moreover, the evidence is *criterial*, not inductive, not deductive, nor *a priori* probabilifying evidence. This point is *absolutely crucial*." (Baker and Hacker, 1980, p. 615.)

Appendix A

The problem of 'visual space' arises when psychological criteria are sought for congruence of shape. A psychological answer may be expected to the question: *when do these shapes count as the same*?, in discussion of 'apparent shape' or 'phenomenal shape'. Perhaps the shapes that count as the same according to psychological criteria will not be the ones that count as the same under physical operations of translation and rotation in space. Perhaps there will be a geometry of apparent shape or phenomenal

shape that is different from the geometry of those physical operations, just as different geometries reflect different criteria for the congruence of shape. (Or else this search for psychological criteria is misconceived.) Once the congruence (i.e., the equivalence) of apparent shape or phenomenal shape is thought to be independent of physical operations, one may search for other criteria, specifically psychological criteria. This is called investigation of the geometry of visual space. Many geometries have been proposed as the geometry of visual space; their number and variety is a catalogue of geometries. Their proliferation shows how psychological criteria fail to constrain the equivalence of apparent shape, or of any shape. There is not a space that is specifically visual, or mental, or of the imagination. "The 'mental' is not *inside* anything, even though much that is, crudely speaking, mental can be concealed; and the 'physical' is not *outside* the mind, since the mind is not a space." (Hacker, 1987, p. 47.)

While particular applications of geometry to vision can be definite and falsifiable, the program of establishing a natural geometry of vision may be extended indefinitely. Once it has been shown that simple and well-known geometries are unlikely designates for the geometry of vision, odd lesser-known geometries may be proposed instead. However, the program of establishing a natural geometry becomes less plausible as it is extended in this way. Such geometries are indeed unfamiliar, and they need to be introduced before their application is explained: this is an obstacle in rhetoric. Moreover, such geometries are regarded as special or unusual with reason. Many are special in that they were developed through investigation of logical consistency or mathematical completeness – they did not emerge from the practical activity of measurement. But then, the notion of a geometry of vision is not sustained by a practice of measurement either; rather, efforts at measurement are made to satisfy dictates of dogma about vision. Unusual geometries (including non-Riemannian geometries, but not yet non-Archimedean geometries to my knowledge) are applied in a monolithic way, to force consistency upon various problems in the study of vision. One purpose of showing that basic geometric invariants are not seen uniformly and for themselves (cf. Niall and Macnamara, 1990) is to show that an entire class of geometries (the Kleinian geometries) provides no viable candidate for a geometry of vision. This class includes many familiar and applied geometries. Then if the program of establishing a natural geometry for vision is to be continued, the candidate geometries that remain will be either esoteric or difficult to apply. The rhetorical force of this argument should not be underestimated. Suppes (1991, p. 48) discusses the results by Foley (1972), and notes that their interpretation constrains the notion of visual space to "lie outside of any of the Riemannian spaces of constant curvature". Suppes continues: "The results however, are disastrous for any simple geometrical theory of visual space, for it requires us to move outside the framework of the standard elementary homogeneous spaces". And thus the proponents of natural geometry will bear a new burden of explanation: they will need to say how their enterprise should be at all illuminating or useful. "I have made the point on several occasions that it may be the case that classical geometry is the wrong model for visual space... we have not yet been successful in finding general principles of visual perception that have the appropriate invariance properties. It is in fact an open question whether satisfactory general principles exist..." (Suppes, 1991, p. 51, and see Suppes, 1995).

Appendix B

Though the ease with which a millstone can be turned depends on its size and shape, the simple magnitude of angular difference does not depend on size or shape. That is, an operation of rotation on a shape can be considered independent of the complexity of that shape. Are response times similar in the Shepard–Metzler effect, in that they are independent of the complexity of stimulus shapes? Cooper and Podgorny (1976, p. 504) claim that a dependence of response times on both angular difference and complexity of shape is a mark of a certain type of explanation for the mental rotation effect: an explanation that postulates operations on features or parts of shapes rather than whole shapes. Cooper and Podgorny (1976, p. 509) find no evidence of a relation between response time and complexity of shape, or between error rate and complexity of shape. They note that their conclusions are "appropriate only for the definition and range of complexity employed in the present experiment" (1976, p. 510). What might it mean that their conclusions should not be generalized beyond a definition of complexity? One possibility is that their failure to find an effect of complexity on response time is a consequence of the application of an unsuitable measure of shape complexity in the experiment. After all, Shepard and Metzler (in Shepard and Cooper, 1982, p. 39) did find differences in the slope of the response time function among three-dimensional objects in their effect. Yet it is difficult to interpret isolated differences in slope that could be associated with differences in shape. Conceivably, this difference could be confounded with effects such as 'familiarity' or 'naturalness'; a difference could be due to a short response time elicited by a familiar or natural view of a complex object (cf. Cooper, 1975, p. 22). An effect of shape complexity on response times would be demonstrated more convincingly if a measure of shape complexity were used to predict patterns of response times. Cooper (1975) and Cooper and Podgorny (1976) used Attneave and Arnoult's (1956, Method I) method to determine a value for the complexity of a variety of non-convex polygons. Cooper (1975) varied the number of vertices in such polygons. She found that the rank order of the slopes and the rank order of the intercepts of their associated 'mental rotation' response time functions did not match the vertices of these polygons in relative numbers. This is taken as evidence "for the failure of complexity variations in the random forms to produce systematic RT differences" (1975, p. 30). (One may agree, or else surmise that the wrong measure of complexity has been applied.). Note Cooper's standard of proof: she seeks systematic differences that are gauged by an explicit measure of shape. Oddly enough, the demand for a metric of shape leaves this theory prey to accounts that postulate single 'salient landmarks' for comparisons, or feature by feature comparisons of shape. Hochberg and Gellman (1977, p. 25) note that "shapes with salient landmarks... have lower slopes and intercepts in their time/angle function than the shapes in which these features have been made less distinguishable...". Then they claim that this effect of the saliency of shape is not an effect of shape complexity, and they cite as evidence Cooper's (1975) failure to find significant effects of shape complexity. A similar standard of proof to Cooper's standard emerges in an article by S. Shepard and D. Metzler (1988, p. 10), who say:

> "...we do not yet have a satisfactory measure of complexity that reflects its perceptual effects (as opposed to the surface features of the physical stimulus) and

that applies equally to what we have called two-dimensional and three-dimensional stimuli. In the absence of such a measure, we cannot make a definitive determination of the extent to which estimated rates of mental rotation for objects differing in dimensionality are determined by dimensionality or by possible differences between these two types of stimuli in psychological complexity.''

The mental rotation hypothesis does not predict that there are shape-specific contributions to response time, since the magnitude of a rotation (that by which the shapes are supposed to be compared) is independent of the internal geometry of a shape in rotation. Just and Carpenter (1985, pp. 166–167) consider this independence to be an advantage of the hypothesis of mental rotation: ''Mental rotation allows subjects to compare the structure of two objects in considerable detail without completely understanding the structure of either one.'' Some effects of shape on response time have been found by experiment (eg. Yuille and Steiger, 1982), but these have been interpreted as effects of complexity, of dimensionality, or of similarity. Now descriptions of complexity, dimensionality, and similarity are not equivalent to a description of shape, because all three – complexity, dimensionality, and similarity – can be predicated of shape. If there is an effect of shape on response time, then the effects of complexity, dimensionality, and similarity may qualify that effect of shape, but they do not explain it. Initially it may seem difficult to demonstrate an effect of shape. A tough standard has been established for the demonstration of such an effect: that a metric or at least an ordering should be given to predict response times. Far from providing a theory of shape perception, the theory of mental rotation has come to rely more and more on the promise of a future theory of shape perception. That is to say, the theory of mental rotation now demands a metric of shape (cf. Hall and Friedman, 1994) which the operation of rotation cannot provide. Discussion over the role of complexity has not abated: the reader is referred to Bethell-Fox and Shepard (1988) and Folk and Luce (1987) for comprehensive discussions.

Table 1

Degrees	Mean RT (sec)	Mean RT (correct)	Mean ln RT	overall d'
Rotation type: single; Axis: x-axis; All observers				
12.5°	0.88	0.88	−0.21	3.56
25°	0.87	0.86	−0.23	3.44
37.5°	1.07	1.05	−0.09	2.90
50°	1.11	1.08	−0.02	2.64
62.5°	1.17	1.18	0.03	2.47
75°	1.36	1.36	0.18	1.97
Rotation type: tandem; Axis: x-axis; All observers				
12.5°	1.04	1.00	−0.14	3.57
25°	1.21	1.18	−0.07	3.22
37.5°	1.34	1.31	0.05	2.66
50°	1.44	1.31	0.14	2.28
62.5°	1.57	1.59	0.20	2.02
75°	1.65	1.68	0.27	1.60

Appendix C

Table 1 displays mean response times and a measure of sensitivity by condition of rotation type (single or tandem) for the second experiment. Mean response time (each of 320 trials) in seconds is shown for trials on which 'same' was the correct response, as is the mean response time when incorrect trials (on which observers responded 'different') are eliminated. The mean of the logarithm of response time (each of 320 trials) is also displayed for trials on which 'same' was the correct response. An overall measure of sensitivity, d', is calculated from the percentage number of times that observers responded 'same', when 'same' was the correct response, and the percentage number of times the observers responded 'same' when 'different' was the correct response (each of 320 trials).

Appendix D

Table 2 displays response times and errors by conditions of observer (KN or VR), rotation type (single or tandem), and axis of rotation (x-axis: pictured rotation in depth, or z-axis: rotation in the picture plane) for the third experiment. Mean response time in seconds is displayed, as is the mean of the logarithm of those response times. Each mean is calculated from 640 observations. Also reported are the percentage of trials (of 320 trials) on which the observer responded 'same' correctly ('hit rate'), and the percentage of trials on which the observer responded 'same' when the two shapes were different ('false alarm rate'). A measure of discriminability, d', can be calculated from these two percentages, following an independent-observation model for same/different comparisons (as in Macmillan and Creelman, 1991).

Appendix E

(Mathematical derivation by Peter Tikuisis)

We are interested in response time differences associated with the rotation of a pair of planar shapes, and the generalization of those differences. The pair of shapes consists of a standard shape and a comparison shape. Angle ψ represents the depicted orientation of the standard shape of a pair, with respect to the picture plane (It represents the initial or reference orientation). Angle θ represents the depicted orientation of the comparison shape of the pair, with respect to the picture plane. We assume that response time differences follow a piecewise continuous function that is defined initially at $\psi = 0$ over a half cycle of θ, and then extended to a full cycle of θ. The angle between the two shapes is given by $\theta - \psi$. (Our concern is with the magnitude of this difference, not with its sign.) We wish to know the form of a response time function for $\theta - \psi$, when ψ is allowed to vary over a half cycle. The form of the piecewise continuous function is defined for $\psi = 0$ as: $-\cos \theta$ from $\theta = 0$ to $\theta = \pi/2$ radians, and $1 + \cos \theta$ from $\theta = \pi/2$ to $\theta = \pi$ radians (see Fig. 20). This function can be extended over a cycle of

Table 2

Degrees	Mean RT (sec)	Mean ln RT	% ('same'/same)	% ('same'/diff.)
Rotation type: single; Axis: x-axis; Observer KN				
15°	0.88	−0.16	100	0
30°	0.96	−0.09	99	0
45°	1.08	0.01	94	2
60°	1.27	0.15	89	3
75°	1.42	0.28	83	8
105°	2.20	0.69	79	24
120°	2.00	0.58	86	17
135°	1.85	0.50	92	11
150°	1.74	0.44	91	8
165°	1.67	0.41	94	5
Rotation type: single; Axis: z-axis; Observer KN				
15°	1.00	−0.06	98	3
30°	1.06	0.00	99	6
45°	1.12	0.04	98	6
60°	1.18	0.09	97	4
75°	1.23	0.14	98	1
105°	1.36	0.23	98	2
120°	1.43	0.26	97	1
135°	1.49	0.31	98	4
150°	1.53	0.33	96	6
165°	1.60	0.36	97	5
Rotation type: tandem; Axis: x-axis; Observer KN				
15°	0.92	−0.13	100	0
30°	1.01	−0.06	99	1
45°	1.12	0.04	93	4
60°	1.27	0.16	89	8
75°	1.51	0.32	82	10
105°	1.97	0.54	92	35
120°	1.67	0.39	93	28
135°	1.54	0.29	95	13
150°	1.34	0.18	95	5
165°	1.07	0.00	98	1
Rotation type: tandem; Axis: z-axis; Observer KN				
15°	1.00	−0.06	98	2
30°	1.05	−0.02	99	4
45°	1.08	0.01	98	1
60°	1.12	0.05	97	0
75°	1.15	0.07	97	0
105°	1.26	0.15	99	0
120°	1.19	0.11	98	1
135°	1.20	0.11	97	1
150°	1.16	0.08	98	1
165°	1.10	0.04	99	1

Table 2 (continued)

Degrees	Mean RT (sec)	Mean ln RT	% ('same'/same)	% ('same'/diff.)
Rotation type: single; Axis: x-axis; Observer VR				
15°	1.20	0.01	98	0
30°	1.25	0.08	98	0
45°	1.52	0.22	99	4
60°	1.79	0.36	97	7
75°	2.21	0.60	92	21
105°	2.24	0.64	85	32
120°	2.05	0.50	93	13
135°	1.90	0.41	94	4
150°	1.67	0.33	97	3
165°	1.56	0.29	98	1
Rotation type: single; Axis: z-axis; Observer VR				
15°	0.95	−0.11	100	0
30°	0.98	−0.08	100	1
45°	1.00	−0.06	100	1
60°	1.06	−0.01	100	1
75°	1.07	0.00	100	1
105°	1.13	0.03	100	0
120°	1.15	0.04	99	1
135°	1.14	0.05	99	1
150°	1.14	0.05	98	0
165°	1.12	0.04	99	2
Rotation type: tandem; Axis: x-axis; Observer VR				
15°	1.16	0.01	96	0
30°	1.39	0.14	98	2
45°	1.67	0.29	98	12
60°	2.00	0.47	98	33
75°	1.88	0.43	96	91
105°	1.51	0.18	99	99
120°	1.77	0.35	99	78
135°	1.72	0.31	99	36
150°	1.51	0.17	99	7
165°	1.16	0.01	98	0
Rotation type: tandem; Axis: z-axis; Observer VR				
15°	0.98	−0.10	99	0
30°	0.93	−0.13	99	1
45°	0.96	−0.11	98	0
60°	0.96	−0.11	99	1
75°	1.02	−0.06	100	0
105°	0.99	−0.08	100	0
120°	1.01	−0.07	99	0
135°	0.96	−0.10	99	1
150°	0.98	−0.09	100	0
165°	0.97	−0.11	99	0

θ: it is $-\cos\theta$ from $\theta = -\pi/2$ to $\theta = \pi/2$ radians, and $1 + \cos\theta$ from $\theta = \pi/2$ to $\theta = 3\pi/2$ radians (compare Fig. 18). Call these parts of the function f and g, respectively. Then what is the form of the response time function for many standard shapes at once, that are depicted to stand in orientation anywhere from $\psi = 0$ to $\psi = \pi$? To make the problem easier, we divide these many shapes into two collections: those shapes that stand between $\psi = 0$ and $\psi = \pi/2$ radians from the picture plane, and those shapes that stand between $\psi = \pi$ and $\psi = \pi/2$ from the picture plane. Or better, we may say there are two collections of standard shapes: those that stand between 0 and $+ \pi/2$ radians away from $\psi = 0$ ($\psi = 0$ is in the picture plane) and those that stand between 0 and $-\pi/2$ radians away from $\psi = \pi$ ($\psi = \pi$ is in the picture plane). We do this, because the two collections are symmetric with respect to the picture plane. Call the family of response time functions associated with the first collection *one*, and call the family of response time functions associated with the second collection *two*. Recall that each response time function has two parts. The first parts f of collection *one* we call f_1, and the first parts f of collection *two* we call f_2. The second parts g of collection *one* we call g_1, and the second parts g of collection *two* we call g_2. One function of f_1 at $\psi = 0$ is $-\cos\theta$, between $\theta = -\pi/2$ and $\theta = \pi/2$. The corresponding function of f_2 at $\psi = \pi$ is $\cos\theta$, because $-\cos(\pi - \theta) = \cos\theta$. Similarly, one function of g_1 at $\psi = 0$ is $1 + \cos\theta$, between $\theta = \pi/2$ and $\theta = 3\pi/2$. Then the corresponding function of g_2 at $\psi = \pi$ is $1 - \cos\theta$, between $\theta = -\pi/2$ and $\theta = \pi/2$, because $1 + \cos(\pi - \theta) = 1 - \cos\theta$. (Two other trigonometric identities will be useful in what follows: $\sin(\pi - \theta) = \sin\theta$, and $\sin(\theta - \pi) = -\sin\theta$). The collections of parts f_1 and g_2 fall in one range, and the collections of parts f_2 and g_1 fall in another range. To find the form of the response time function for $\theta - \psi$, we find the normalized mean value of $f_1 + g_2$ within their range, and the normalized mean value of $f_2 + g_1$ within their range. Given that we have these names for the families of curves, we proceed to find the form of the overall function, as follows:

A member of the family of curves represented by f_1 is given by:

$$-\cos(\theta - \psi), \quad \text{where } 0 \le \psi \le (\theta + \pi/2).$$

We wish to integrate this function over the interval.

The mean value of the family of curves represented by f_1 at any value of θ is given by integrating the above function over the interval defined for ψ, and through normalization by ψ over the same interval:

$$\overline{f_1} = \left(\int_0^{\theta + \pi/2} -\cos(\theta - \psi)\mathrm{d}\psi \right) \Big/ \left(\int_0^{\theta + \pi/2} \mathrm{d}\psi \right),$$

where the denominator is the normalization factor. Then

$$\overline{f_1} = -((1 + \sin\theta)/(\theta + \pi/2)).$$

By consideration of symmetry,

$$\overline{f_2} = \overline{f_1}(\pi - \theta),$$
$$\overline{f_2} = -((1 + \sin\theta)/(3\pi/2 - \theta)).$$

Similarly,

$$\overline{g_1} = 1 + \overline{f_1}(\theta - \pi),$$

$$\overline{g_1} = 1 - ((1 - \sin\theta)/(\theta - \pi/2)).$$

And finally,

$$\overline{g_2} = \overline{g_1}(\pi - \theta),$$

$$\overline{g_2} = 1 - ((1 - \sin\theta)/(\pi/2 - \theta)).$$

The overall function combines $(\overline{f_1} + \overline{g_2})$ and $(\overline{f_2} + \overline{g_1})$ in the regions $-\pi/2 \le \theta \le \pi/2$ and $\pi/2 \le \theta \le 3\pi/2$, respectively; note that the final combined form must be re-normalized.

$$\overline{f_1 + g_2} = \frac{(-(1 + \sin\theta) + \pi/2 - \theta - (1 - \sin\theta))}{(\theta + \pi/2 + \pi/2 - \theta)}$$

$$= 1/2 - ((\theta + 2)/\pi), \qquad -\pi/2 \le \theta \le \pi/2.$$

$$\overline{f_2 + g_1} = \frac{(-(1 + \sin\theta) + \theta - \pi/2 - (1 - \sin\theta))}{(3\pi/2 - \theta + \theta - \pi/2)}$$

$$= -1/2 + ((\theta - 2)/\pi), \qquad \pi/2 \le \theta \le 3\pi/2.$$

Together these define a *triangle function*, that has the piecewise linear form that is characteristic of response time differences for the Shepard–Metzler effect in depth with very complex shapes.

The triangle function $T(\theta)$ can be given the following conditional form:

If $\quad -\pi/2 \le \theta \le \pi/2$, then $\quad T(\theta) = 1/2 - ((\theta + 2)/\pi)$.   (1)

If $\quad \pi/2 \le \theta \le 3\pi/2$, then $\quad T(\theta) = -1/2 + ((\theta - 2)/\pi)$.   (2)

If $\quad \theta > 3\pi/2$, then reset $\quad \theta = \theta - 2\pi$.   (3)

(*Repeat until either* $(-\pi/2 \le \theta \le \pi/2)$ *or* $(\pi/2 \le \theta \le 3\pi/2)$.)

If $\quad \theta < \pi/2$, then reset $\quad \theta = \theta + 2\pi$.   (4)

(*Repeat until either* $(-\pi/2 \le \theta \le \pi/2)$ *or* $(\pi/2 \le \theta \le 3\pi/2)$.)

References

Attneave, F. and M.D. Arnoult, 1956. The quantitative study of shape and pattern perception. Psychological Bulletin 53(6), 452–471.

Auden, W.H., 1966. 'In praise of limestone'. Collected shorter poems 1927–1957. London: Faber and Faber.

Baker, G.P. and P.M.S. Hacker, 1980. Wittgenstein: Understanding and meaning. Oxford: Basil Blackwell.

Bauer, B. and P. Jolicoeur, 1996. Stimulus dimensionality effects in mental rotation. Journal of Experimental Psychology: Human Perception and Performance 22(1), 82–94.

Bethell-Fox, C.E. and R.N. Shepard, 1988. Mental rotation: Effects of stimulus complexity and familiarity. Journal of Experimental Psychology: Human Perception and Performance 14(1), 12–23.

Boring, E.G., 1952. Visual perception as invariance. Psychological Review 59, 141–148.

Carlton, E.H. and R.N. Shepard, 1990. Psychologically simple motions as geodesic paths: I. Asymmetric objects. Journal of Mathematical Psychology 34, 127–188.

Cohen, D. and M. Kubovy, 1993. Mental rotation, mental representation, and flat slopes. Cognitive Psychology 25, 351–382.

Cooper, L.A., 1975. Mental rotation of random two-dimensional shapes. Cognitive Psychology 7, 20–43.

Cooper, L.A. and P. Podgorny, 1976. Mental transformations and visual comparison processes: Effects of complexity and similarity. Journal of Experimental Psychology: Human Perception and Performance 2(4), 503–514.

Cooper, L.A. and R.N. Shepard, 1973. 'Chronometric studies of the rotation of mental images'. In: W.G. Chase (Ed.), Visual information processing. New York: Academic Press, pp. 75–176.

Cooper, L.A. and R.N. Shepard, 1984. Turning something over in the mind. Scientific American 251(6), 106–114.

Corballis, M.C., 1988. Recognition of disoriented shapes. Psychological Review 95, 115–123.

Farah, M.J., R. Rochlin and K.L. Klein, 1994. Orientation invariance and geometric primitives in shape recognition. Cognitive Science 18, 325–344.

Foley, J.M., 1972. The size–distance relation and intrinsic geometry of visual space: Implications for processing. Vision Research 13, 323–332.

Folk, M.D. and R.D. Luce, 1987. Effects of stimulus complexity on mental rotation rate of polygons. Journal of Experimental Psychology: Human Perception and Performance 13(3), 395–404.

Foster, D.H. and D.R. Simmons, 1994. Viewpoint-invariance of contour curvature discrimination over extended performance levels. Spatial Vision 8(1), 45–55.

Greenhouse, S.W. and S. Geisser, 1959. On methods in the analysis of profile data. Psychometrika 24(2), 95–112.

Hacker, P.M.S., 1987. Appearance and reality. Oxford: Basil Blackwell.

Hacker, P.M.S., 1990. Wittgenstein: Meaning and mind. Oxford: Basil Blackwell.

Hall, D.L. and A. Friedman, 1994. Shape discriminations of three-dimensional objects depend on the number and location of bends. Perception and Psychophysics 56(3), 288–300.

Hochberg, J. and L. Gellman, 1977. The effect of landmark features on mental rotation times. Memory and Cognition 5(1), 23–26.

Jolicoeur, P., 1990. Identification of disoriented objects: A dual-systems theory. Mind and Language 5(4), 387–410.

Jolicoeur, P., S. Regehr, L.B.J.P. Smith and G.N. Smith, 1985. Mental rotation of representations of two-dimensional and three-dimensional objects. Canadian Journal of Psychology 39(1), 100–129.

Just, M.A. and P.A. Carpenter, 1985. Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. Psychological Review 92(2), 137–171.

Koriat, A. and J. Norman, 1985. Mental rotation and visual familiarity. Perception and Psychophysics 37, 429–439.

Lappin, J.S. and U.B. Ahlström, 1994. On the scaling of visual space from motion – In response to Pizlo and Salach-Golyska. Perception and Psychophysics 55(2), 235–242.

Leone, G., M.-C. Taine and J. Droulez, 1993. The influence of long-term practice on mental rotation of 3-D objects. Cognitive Brain Research 1, 241–255.

Macmillan, N.A. and C.D. Creelman, 1991. Detection theory: A user's guide. New York: Cambridge University Press.

Massaro, D.W., 1973. The perception of rotated shapes: A process analysis of shape constancy. Perception and Psychophysics 13(3), 413–422.

Metzler, J. and R.N. Shepard, 1974. 'Transformational studies of the internal representation of three-dimensional objects'. In: R.L. Solso (Ed.), Theories of cognitive psychology: The Loyola symposium. Potomac, MD: Erlbaum, pp. 147–201.

Niall, K.K. and J. Macnamara, 1990. Projective invariance and picture perception. Perception 19, 637–660.

Parsons, L.M., 1987. Visual discrimination of abstract mirror-reflected three-dimensional objects at many orientations. Perception and Psychophysics 42(1), 49–59.

Parsons, L.M., 1987a. 'Spatial transformations used in imagination, perception, and action'. In: L.M. Vaina (Ed.), Matters of intelligence. Dordrecht: Reidel, pp. 143–181.

Parsons, L.M., 1995. Inability to reason about an object's orientation using an axis and angle of rotation. Journal of Experimental Psychology: Human Perception and Performance 21(6), 1259–1277.

Petrusic, W.M., L. Varro and D.G. Jamieson, 1978. Mental rotation validation of two spatial ability tests. Psychological Research 40, 139–148.

Putnam, H., 1990. Realism with a human face. Cambridge, MA: Harvard University Press.

Pylyshyn, Z.W., 1981. 'The imagery debate: Analog media versus tacit knowledge'. In: N. Block (Ed.), Imagery. Cambridge, MA: MIT Press, pp. 151–206.

Ratcliff, R., 1993. Methods for dealing with reaction time outliers. Psychological Bulletin 114(3), 510–532.

Rock, I., D. Wheeler and L. Tudor, 1989. *Can* we imagine how objects look from other viewpoints? Cognitive Psychology 21, 185–210.

Shepard, R.N., 1978. The mental image. American Psychologist 33, 125–137.

Shepard, R.N., 1987. 'The role of transformations in spatial cognition'. In: J. Stiles-Davis, M. Kritchevsky and U. Bellugi (Eds.), Spatial cognition: Brain bases and development. Hillsdale, NJ: Erlbaum, pp. 81–110.

Shepard, R.N. and L.A. Cooper, 1982. Mental images and their transformations. Cambridge, MA: MIT Press.

Shepard, R.N. and S.A. Judd, 1976. Perceptual illusion of rotation of three-dimensional objects. Science 191, 952–954.

Shepard, R.N. and J. Metzler, 1971. Mental rotation of three-dimensional objects. Science 171, 701–703.

Shepard, S. and D. Metzler, 1988. Mental rotation: effects of dimensionality of objects and type of task. Journal of Experimental Psychology: Human Perception and Performance 14(1), 3–11.

Suppes, P., 1991. 'The principle of invariance with special reference to perception'. In: J.-P. Doignon and J.-C. Falmagne, Mathematical psychology: current developments. New York: Springer-Verlag, pp. 35–53.

Suppes, P., 1995. 'Some foundational problems in the theory of visual space'. In: R.D. Luce et al. (Eds.), Geometric representations of perceptual phenomena. Mahwah, NJ: Erlbaum.

Van Gool, L.J., T. Moons, E. Pauwels and J. Wagemans, 1994. Invariance from the Euclidean geometer's perspective. Perception 23, 547–561.

White, M.J., 1980. Naming and categorization of tilted alphanumeric characters do not require mental rotation. Bulletin of the Psychonomic Society 15(3), 153–156.

Wittgenstein, L., 1953. Philosophical investigations. Oxford: Basil Blackwell (G.E.M. Anscombe, trans.).

Yuille, J.C. and J.H. Steiger, 1982. Nonholistic processing in mental rotation: some suggestive evidence. Perception and Psychophysics 31(3), 201–209.